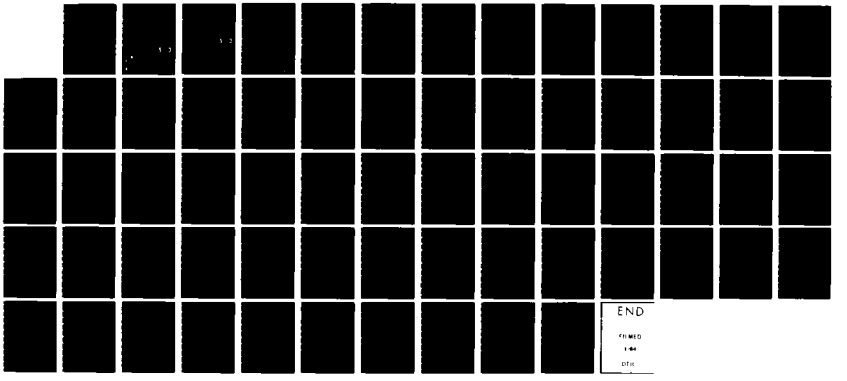
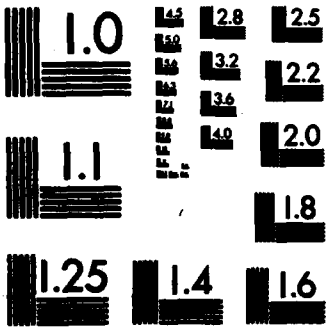
AD-A135 501 TECHNIQUES OF PSYCHOSOCIAL EPIDEMIOLOGY(U) NAVAL HEALTH 1/1
RESEARCH CENTER SAN DIEGO CA L A PALINKAS OCT 83
NAVHLTHRSCHC-83-22

UNCLASSIFIED F/G 6/5 NL





(2)

TECHNIQUES OF PSYCHOSOCIAL EPIDEMIOLOGY

L. A. PALINKAS

REPORT NO. 83-22

NAVAL HEALTH RESEARCH CENTER
P. O. BOX 85122
SAN DIEGO, CALIFORNIA 92138

NAVAL MEDICAL RESEARCH AND DEVELOPMENT COMMAND
BETHESDA, MARYLAND

TECHNIQUES OF PSYCHOSOCIAL EPIDEMIOLOGY

Lawrence A. Palinkas, Ph.D.

Research Associate, National Research Council

Naval Health Research Center

P.O. Box 85122

San Diego, California 92138-9174

To expedite communication of the research, this is a preprint of a chapter in the volume, Measurement Strategies in Health Psychology, edited by Paul Karoly, and should be cited as a personal communication.

Report Number 83-22

supported by the Naval Medical Research and Development Command, Department of the Navy, under Research Work Unit MR0000.01.01-6033 and the National Research Council. The views presented in this paper are those of the author. No endorsement by the Department of the Navy has been given nor should any be inferred.

Summary

Problem

As new information is obtained about the relationships among the physiological, social and environmental aspects of disease, the methods employed by epidemiologists have been revised and broadened. One example of this refining process is the emergence of the distinct subfield of psychosocial epidemiology. While exhibiting a certain measure of continuity with other areas of epidemiologic research in its approach to disease and methodology employed, psychosocial epidemiology is characterized by its selective focus on the psychosocial aspects of disease and its reliance on psychometric methods. Because of its interdisciplinary approach, a clarification of the techniques employed in psychosocial epidemiology would be extremely useful to epidemiologists, health psychologists, and other medical researchers.

Objectives

The objective of the chapter is to summarize the techniques employed in psychosocial epidemiology by providing a context for their use and demonstrating the ways these techniques are based on the particular focus of investigation. In addition, the strengths and weaknesses of these techniques are examined.

Approach

The approach to this review is to begin with an outline of the epidemiologic foundations of the concepts and methods of psychosocial epidemiology and to provide some understanding of what distinguishes the psychosocial approach to other forms of epidemiologic research. The review then proceeds with a discussion of the types of research design, data, and data analysis employed in psychosocial studies. The chapter concludes with a brief discussion of the importance of qualitative techniques to psychosocial epidemiology.

Results

Psychosocial epidemiology embodies a range of studies on a biologic-social-psychological continuum, incorporating certain aspects of traditional biomedical epidemiology and psychiatric epidemiology. It is distinguished from these other two forms of epidemiology by its reliance on psychometric methods and its emphasis on the role of stress in the distribution and spread of disease.

The selection of an appropriate research design is based on a series of considerations regarding the research objectives and available resources. The types of research design used by psychosocial epidemiologists can be arranged in a taxonomic fashion, with observational and experimental studies at the broadest level. Observational methods are most commonly employed and study variables are indirectly controlled through the use of quantitative methods. Two types of observational studies are descriptive and analytic studies. These, in turn, are divided into cross-sectional, retrospective and prospective studies. These are distinguished from one another on the basis of time perspective and direction of causality. In the cross-sectional study, the characteristic or type of exposure being compared is present in both cases (those having experienced a disease or stress event) and

controls at the time of the investigation. In the retrospective study, the object is to determine whether the characteristic was present in the past. In the prospective study, a cohort is selected and subdivided on the basis of characteristic or exposure. The researcher then follows the cohort for an extended period of time to determine which members experience the disease event and if a particular pattern emerges with respect to one or more characteristics.

The data for psychosocial studies are usually represented in the form of rates and scales. Each form employs a characteristic set of methods for the organization and expression of dependent and independent variables. The major dependent variables in psychosocial studies include disease and stress events. Disease events are expressed in terms of mortality and morbidity rates. Incidence and prevalence rates are the major forms of morbidity rate, distinguished by whether the number of events observed over a period of time are new cases or all cases observed throughout the period. Stress events are measured by psychosocial scales which can be subdivided into the categories of self-report, performance, psychophysiological and biochemical measurement scales. The major independent variables in psychosocial studies include sociodemographic characteristics, life events, social support indices, and personality measures. The limitations of each of these forms of data, including problems of definition, sources of data, weaknesses of stress measurements, and the difficulties of operationalizing certain independent variables, are discussed.

Data analysis proceeds with the execution of three discrete steps: description of the phenomena under investigation, search for associations, and multivariate analyses. The measures of association described in this chapter include relative risk, chi-square tests, correlation coefficients, lifetime risk and attributable risk. Methods of association used specifically in prospective studies include life tables, survival analysis, and Fourier analysis. Multivariate analyses are used to adjust and control for potential confounding variables. Variable-specific rates and direct and indirect age-adjustment procedures are presented as means to control for such variables. The principles of regression analyses are briefly reviewed as an introduction to the application of logistic regression as a means of controlling for confounding variables and hypothesis testing.

The last section of the chapter briefly touches upon the necessity for a qualitative perspective in each of the stages of a psychosocial study: study design, data collection, and data analysis. Methods such as participant observation, informal interviews, and thematic analysis are presented as techniques useful for detecting potential characteristics for study, verifying the quality of the data collected, and avoiding spurious relationships in analysis.

Techniques of Psychosocial Epidemiology

INTRODUCTION

Epidemiology is concerned with the origin and spread of disease in human populations. In its infancy, epidemiology was primarily concerned with infectious diseases; later, chronic disease became a focus for intensive research by epidemiologists. As new information is obtained about the relationships among the physiological, social and environmental aspects of disease, the methods employed by epidemiologists have been revised and broadened. One example of this refining process is the emergence of a distinct subfield of psychosocial epidemiology. Psychosocial epidemiology displays a certain measure of continuity with other areas of epidemiological research in its approach to disease and methodology employed. It diverges from other epidemiologic studies in its selective focus on psychosocial components of disease and greater reliance on psychometric methods. Psychosocial epidemiology has often been identified with an interest in the problem of stress, which serves to link the psychological, social, environmental, and physiological components of health.

This chapter is intended to summarize the techniques employed in psychosocial epidemiology. The summary is divided into three major sections: research design, data, and data analysis. Many of these techniques will be discussed in greater detail elsewhere in the volume; the objective here is to provide a context for their use and to demonstrate the ways these techniques are based on the particular focus of investigation. In the course of this presentation, the strengths and limitations of these techniques will be examined and evaluated.

Epidemiologic Foundations

Before detailing the concepts and methods employed in psychosocial epidemiology, a brief overview of the epidemiologic foundation of these concepts and methods is necessary. Epidemiology can best be defined as "the study of the distribution of a disease or a physiological condition in human populations and of the factors that influence this distribution" (Lilienfeld, 1978, p. 87). Although epidemiology utilizes qualitative, clinical data, the emphasis is on quantitative data and analyses. According to Lilienfeld and Lilienfeld (1980), the epidemiologist uses a two-stage sequence of reasoning. The first stage involves the determination of a statistical association between a characteristic and a disease. The second stage is the derivation of biological inferences from such a pattern of statistical inferences.

The distribution of a disease in a particular population may be influenced by one or more of three sets of variables: agent, host, and environment. The agents of disease include etiological factors including nutritive elements such as vitamins or minerals, chemical agents such as poisons or allergens, or physical agents such as ionizing radiation and noise. A second set of variables pertain to the disease hosts, the individual or individuals experiencing the disease episode. These include genetic factors, demographic

characteristics such as age, sex, and race, social factors such as education and income, cultural factors such as ethnic group and religion, behavioral factors such as stress-coping strategies, health-impairing habits and lifestyles, and personality factors such as ego defense mechanisms and self-esteem. The third set of variables comprise the environment in which a disease episode occurs. This includes the physical environment (geology, climate), biological environment (flora, fauna), and social environment (social networks, residence, occupation). These three sets of variables together form the epidemiologic triad of agent, host and environment and their interrelations are the primary focus of investigation.

Psychosocial Epidemiology

Although psychosocial epidemiology is usually identified with a cluster of studies employing particular screening scales or survey methods, this subdiscipline can actually be said to embody a range of epidemiologic studies on a biological-social-psychological continuum. Thus, certain aspects of traditional biomedical epidemiology and of psychiatric epidemiology can be considered part of psychosocial epidemiology. What distinguishes the psychosocial from the biomedical and psychiatric forms of epidemiology is a focus incorporating all three components (biological-social-psychological) and their interrelations. As such it shares certain techniques with the other two branches of epidemiology and employs certain techniques which are unique. Similarly, elements of psychosocial epidemiology can be found in the other two branches.

A distinguishing characteristic of psychosocial epidemiology is the emphasis on the role of stress in the distribution and spread of disease. Although several definitions of stress exist, perhaps the best known is the one provided by Selye (1956) who defined it as a nonspecific response of the body to any demand. Since the pioneering work of Cannon (1929) and Selye (1956), several models have been developed which highlight the importance of physiological and psychological manifestations of stress in the etiology and outcome of disease. The basis of these models is the notion that environmental factors may act to create a physical or psychological strain on the host, be it an individual or a social group. These environmental factors may be present or anticipated, physical, psychological or sociocultural. These stressors serve to stimulate both the pituitary-adrenal axis (Selye, 1956) and the sympathetic-adrenal medullary system (Cannon, 1929), producing increased levels of catecholamines and corticosteroids. While this stimulation is viewed as an adaptive response, sustained levels of stress can result in the weakening of the immunological system and excessive strain on the body in general, leading to increased susceptibility to both infectious and chronic diseases. Despite the need for further elaboration of existing models, the notion of stress is an important one in psychosocial epidemiology, tying together all three components of illness: social, psychological and physiological.

RESEARCH DESIGN

Several considerations are involved in the selection of a particular research design in a psychosocial epidemiologic study. For the sake of convenience, these considerations can be grouped into the categories of objectives and resources. Before undertaking a study, it is important to have a clear sense of what needs to be done. This involves specification of issues underlying the phenomena to be investigated, data needs and data limitations, hypotheses to be tested, and methods available for data analysis. Once a decision has been made as to what is needed to be done and/or what is desirable in investigating the phenomena of interest, these decisions must be evaluated in terms of what can be done given available resources. The characteristics of available resources pose both opportunities and limitations to an epidemiologic study. Before deciding what type of study design to employ in investigating a specific disease or stress response and set of characteristics, the nature of the resources involved in both obtaining information on the phenomena under investigation and analyzing the data must be ascertained. Availability of time, money, and personnel are always important considerations; insufficient quantities of either will limit the options available for data gathering and analysis. The type of data available also plays an important role in the selection of a particular study design.

There are several different study designs available in psychosocial epidemiology and distinguishing among them on the basis of their objectives and methods can be difficult at best. The use of the same label to refer to different types of study or different labels to refer to the same type of study often leads to confusion as to the design being implemented (c.f., Feinstein, 1979). Thus, we will begin with an attempt to summarize the criteria used to classify epidemiologic studies.

Observational and Experimental Studies

At the broadest level, epidemiologic studies can be grouped into two major categories, observational and experimental, distinguished on the basis of the degree of control exercised by the researcher. Observational studies are those in which the phenomena under investigation are not amenable to control by the researcher. The researcher cannot specify the conditions under which a particular disease pattern occurs but must observe the event in its natural context. Experimental studies, on the other hand, are subject to direct intervention by the researcher. Variables can be manipulated and controlled, thus reducing the possibility of spurious relationships and confounding factors. Performance studies of stress in which subjects are exposed to similar experiences to determine their biochemical reactions to stress are examples of such studies. When subjects can be randomly allocated into exposed and nonexposed groups, this study is known as a clinical trial (Lilienfeld and Lilienfeld, 1980). In a clinical trial, the focus of investigation is upon the individual. When random allocation is not possible, the experimental study is known as a "community trial." Studying the effects of changes in treatment policy among military personnel diagnosed as alcoholics is an example of a community trial. The policy is subject to manipulation by the investigators, but random allocation of military personnel into cases and controls is not possible, nor can all extraneous variables be control-

led for. Moreover, the focus of the study is the group as a whole (i.e., military personnel), not individuals within the group.

Descriptive and Analytic Studies

Because of the difficulty of conducting well-controlled epidemiologic experiments on human populations on the one hand and the availability of observations of past or ongoing events on the other, observational studies are more common and the variables under study are indirectly controlled through the use of quantitative methods of data analysis. There are two major categories of observational study, descriptive and analytic. The objective of a descriptive study is to describe the phenomena under investigation through the use of standardized rates of morbidity or mortality, the amount of variability in sets of observations among groups of individuals, and relationships among dependent, independent, and intervening variables of interest. Such studies utilize rates of incidence, prevalence, or mortality in describing the pattern of disease in a population on the basis of time, place, and population characteristics such as age, sex, and race.

Analytic studies are typically conducted to explain the observed pattern of disease occurrence. Their purpose is to determine whether or not an association is present between a certain characteristic or combination of characteristics and the disease or level of stress in a group of afflicted individuals. Whenever possible, the causal relationship between independent and dependent variables is also examined. In these studies, comparisons are made between a group of persons who have the disease under investigation and a control group which does not. There is a great deal of overlap between descriptive and analytic studies but the latter is usually more focussed and concerned with the testing of specific hypotheses.

Both descriptive and analytic studies can be further subdivided on the basis of causal and time perspectives. Cross-sectional and retrospective studies focus on the dependent variable, proceeding from effect to cause. These studies are often referred to as case-control studies, although, as Lilienfeld and Lilienfeld (1980) observe, the case-control perspective is also utilized in prospective studies. Prospective or cohort studies focus on the independent variables, proceeding from cause to effect. All three types of study design involve comparisons among individuals who experience a disease episode and individuals who do not with respect to one or more characteristics or exposures. The model for these comparisons, represented in its simplist form, is a 2x2 table such as Table 1.1

Cross-sectional and Retrospective Studies The object of cross-sectional and retrospective studies is to determine whether or not a relationship exists between a particular trait or set of traits and a specific illness or disease. Subjects are selected on the basis of presence or absence of the disease under investigation and then compared on the basis of one or more characteristics. Subjects with the disease are referred to as cases while those who do not have the disease serve as controls. If a higher proportion of individuals with the characteristic is found among the cases than the controls, an association between the disease and the characteristic is indicated.

Table 1.1

Framework for An Epidemiologic Study

Prospective (Cohort)

Effects

Number of Individuals

Cross-sectional/ Retrospective Causes: Characteristic	With Disease	Without Disease	Total
With	a	b	$a + b = N_1$
Without	c	d	$c + d = N_2$
Total	$a + c = M_1$	$b + d = M_2$	$a + b + c + d = N$

Cross-sectional and retrospective studies are distinguished from one another on the basis of time. In the cross-sectional study, the characteristic being compared is present in both cases and controls at the time of the investigation. These studies are also known as "prevalence" studies. The study of suicidal behavior in basic military trainees by Gaines and Richmond (1980) provides an example of a cross-sectional study. Sixty basic trainees were selected for the study, 30 of whom had been brought to the attention of authorities for having made a suicidal gesture, and 30 non-suicidal trainees who served as a control group. Both cases and controls were evaluated on the same day and each subject was independently administered a questionnaire to obtain information on birth order, parent's ages, education, occupation, marital status, and so on. Perceptions of pretraining and basic training experiences were obtained and each subject was administered the Wechsler Adult Intelligence Scale (WAIS) and the Minnesota Multiphasic Personality Inventory (MMPI), using the standard administrative procedures.

Another example of a case-control study is provided by Harburg, Gleiberman, Roeper, Schork, and Schull (1978) in their examination of the relationships between skin color, ethnicity, stress, and blood pressure in Detroit. Census areas in Detroit were ranked for their stress scores on the basis of their instability (measured in terms of levels of crime, marital breakup and so on) and socioeconomic status. Four areas were selected for detailed study: 1) high stress, population predominantly black; 2) high stress, population predominantly white; 3) low stress, population predominantly black; and 4) low stress, population predominantly white. A random sample of 1,000 adults was drawn from each area of persons of the predominant race in the census division, 25-60 years old, married and living with spouse, and having relatives in the Detroit area. Each subject was interviewed by a nurse who took three blood pressure readings and rated skin color on a four-point scale. Both the dependent variable (blood pressure) and independent variables (race, skin color, residence, etc.) were present in the sample at the time of the study. Results showed that darker skin color, for black males especially, was related to higher blood pressure, independent of control variables such as age, weight, height, smoking history, educational level, and family income.

In a retrospective study, the object is to determine whether the characteristic was

present in the past. The investigator looks backward in time for exposure. Most studies of stressful life events are representative of this approach. In a study of life events and disease by Antonovsky and Kats (1967), a case sample of patients with multiple sclerosis was matched with a control sample on the basis of sex, age and region of birth. The two groups were compared on the basis of "life crises," defined as events associated with physical trauma and changes in environment, primary interpersonal relations, and status. The events were then operationalized to create a "crisis score" to indicate the intensity of the crises.

In the set-up of a case-control design, several considerations must be taken into account. The first and most important is the determination of where the cases and controls will be obtained. Both may be drawn from the community at large. In a study of depression for instance, all diagnosed cases in a community, using hospital admissions, clinic visits, or physicians' records as a source of data, could serve as cases while the controls would consist of a sample of non-depressed individuals from the general population in a community. A second approach would be to select a sample of the general population and divide those diagnosed as being clinically depressed (on the basis of clinical examinations or the administration of a screening scale such as the CES-D) to comprise the cases and those evaluated as not depressed serving as the controls. A third option is to select a specific population, such as the patients at a particular hospital or group of hospitals, diagnosed as being clinically depressed, and compare them with a group of patients with other diagnoses. The former group would serve as the cases while the latter would comprise the controls in the study.

Another consideration in a case-control design involves taking into account known or suspected factors believed to be related to the disease under investigation. The idea behind random sampling in epidemiologic studies is to control in part for extraneous variables which could possibly influence the distribution of the disease but which cannot be anticipated ahead of time. Adjustment procedures can also be implemented to account for potential bias. Cases and controls can be compared for each level of extraneous factor such as age, sex, socioeconomic status, social supports, and so on. Cases and controls can also be matched with one another, either as a group or as individuals. A group match would stratify both the case group and the control group into levels, such as age groups. For example, the cases can be stratified into ten-year age groups, 20-29, 30-39, 40-49, and so on. The control group can be similarly stratified and samples of controls randomly selected from each stratum to provide the same number of controls as cases. If the number of characteristics to be examined and controlled for is not too large, individual matching may be performed, resulting in pairs of cases and controls possessing the same characteristic or limited set of characteristics (e.g., black males, black females white ma' white females).

Several procedures are available to determine how large the sample should be for a particular study. As Schlesselman (1982, p. 144) notes, a study should be large enough to avoid both Type I (claiming an association when one does not exist) and Type II (claiming

no association when one does exist) statistical errors . The number of subjects to be selected for a case-control study depends on the specification of four values: (1) the relative frequency of exposure among controls in the target population, (2) a hypothesized relative risk associated with exposure that would have sufficient biologic or public health importance to warrant its detection, (3) the desired level of significance, and (4) the desired study power. Schlesselman (1982) provides a formula which incorporates these four values in the determination of a satisfactory sample size.

There are several advantages to using a cross-sectional or retrospective design when conducting an epidemiologic study. They are well suited to the study of rare diseases or those with long latency. They are relatively quick to organize and carry out. Compared with the time and effort of a cohort study, they are relatively inexpensive and require comparatively few subjects. Within certain limits, they have the advantage of utilizing existing records and do not involve any undue risk to the subjects. Finally, a cross-sectional or retrospective design allows for the study of multiple potential causes of a disease (Schlesselman, 1982).

On the other hand, there are several disadvantages to the case-control format of cross-sectional and retrospective designs. Retrospective studies, for example, rely on recall or records for information on past exposure and validation of this information is difficult, at times impossible, (Schlesselman, 1982, p. 18). As will be noted when specific types of data are examined below, the great potential for incomplete data and confounding factors in psychosocial studies severely restrict the use of these designs. Even though the case-control procedure in these designs allows for an examination of multiple independent variables, the control of these variables is likely to be incomplete. Finally, relative to the cohort design, cross-sectional and retrospective designs are inadequate as predictors of a disease or stress event. Such a design can only give estimates of relative risk and not a direct estimate of probability because, by looking at a particular segment of cases, those present at the time of the investigation, one cannot ascertain the disease episodes among cases not present. To utilize the characteristic under study in an effort to predict the likelihood of experiencing the disease or stress event, the inference must assume the form: "Given the existence of a charactertistic (e.g., race, sex, life event), what is the probability that an individual will experience a particular event (e.g., psychological distress, depression, hypertension, ulcers)?" Answers to this question can only be obtained from prospective evidence (c.f., Hirschfeld and Cross 1982:42).

Prospective Studies The prospective or cohort study begins with a specified population believed to be free of the stress, disease or diseases under investigation. The population is divided into groups on the basis of one or more characteristics which may affect the risk of experiencing the stress or disease over time. The researcher looks forward in time for exposure and analysis proceeds from cause to effect. In this approach, also known as a follow-up study, the population may be followed for several years to determine which members experience the disease event and if a particular pattern emerges with respect to one or more characteristics. The population is divided into subgroups according

to the presence or absence of one or more characteristics (or the degree to which the characteristic is present or absent) initially. The subgroups are then compared with respect to the subsequent incidence of disease over a specified follow-up period. A higher incidence of disease in a subgroup possessing the characteristic indicates an association between the disease and the characteristic. This method provides the most direct measurement of the risk of disease development and hence is preferred by epidemiologists.

There are two types of prospective study, concurrent and nonconcurrent. Concurrent studies begin with a presently existing cohort which is followed for an extended period of time into the future. Residents of a particular city, students attending a particular college, all enlisted personnel in the U.S. Navy, who are to be observed and followed for the next ten years would be examples of such a study. A nonconcurrent study selects a cohort as it existed at some point in the past (students who graduated from a college in 1970; residents of a city in 1950-1955; recruits who entered the Navy in 1973) and follows them for a designated period of time, often to the present.

An example of a concurrent, prospective, psychosocial study is the 1965 Human Population Laboratory survey of a random sample of adults in Alameda County, California (Berkman and Syme, 1979). A stratified sample of Alameda County housing units was selected and visited by an enumerator who gathered demographic data on all household members of all ages. The cohort was subdivided on the basis of several characteristics, including age, race, sex, socioeconomic status, and social networks. Mortality data was then collected for a nine-year period, from 1965 to 1974, when a follow-up survey was conducted. Death certificates from those belonging to the 1965 cohort were obtained using the computer matching files of the California Death Registry or other state registries when the cohort member had moved out of the area. Through extensive follow-up, all but 4 percent of the original cohort had been located. Those lost to follow-up were not found to differ significantly from those accounted for on any of the health measures recorded in the original survey, thus reducing the possibility of a spurious relationship in survivorship. The mortality experience of particular subgroups were then compared to see if any association existed between one or more characteristics, such as social networks, and mortality risk.

Prospective studies have also been employed in studies of disease and life events. In a study by Grant, Yager, Sweetwood, and Olshen (1982), a cohort consisting of psychiatric outpatients from a Veterans Administration Mental Health Clinic and nonpatients recruited from a university community was selected and observed over a three-year period. At the beginning of the study, all subjects were interviewed and demographic, biographic, event and symptoms data obtained. For the next three years, each member of the cohort was sent a copy of the Holmes and Rahe Schedule of Recent Experiences (SRE) and a Symptom Checklist (SCL) every two months. Subjects indicated on the SRE which events occurred to them in the previous two months and rated all the items on the scale on a seven-point desirability scale, whether or not the events had occurred. Thus, for every two-month interval, it was possible to compute an event score measured by life change units (LCUs). Symptom data

were obtained from a 69-item SCL of bodily, cognitive, affective and behavioral symptoms.

The procedure for conducting a prospective study is relatively straightforward. The cohort selected can be a sample from a general population group such as a community, or a specific group such as the employees of a certain corporation or members of a particular occupation. A cohort may also be selected on the basis of a known exposure to a particular etiologic factor such as individuals who have experienced a divorce or death of spouse or who work in a stressful occupation. In the latter example, an unexposed cohort should also be selected in order to evaluate the experience of the exposed cohort.

While selection of a cohort may be a relatively simple procedure, the follow-up phase of a prospective study is much more difficult, requiring considerable resources. According to Friedman (1980), the length of the follow-up period depends primarily on the number of disease cases needed to provide reliable, statistically significant answers to the specific questions under study. This estimate requires a knowledge of the size of the cohort and an estimate of the incidence rate of the disease episode being examined. This will allow for the estimation of the number of new episodes occurring in a one year period. For example, if there are 8,000 members in the cohort and incidence rate for the disease episode is 1 percent per year, about 80 new cases of the disease episode may be expected during each year of follow-up. If 1600 cases are needed to provide answers with a certain degree of reliability, then the study may be expected to last about 10 years (c.f., Friedman, 1980, p. 124).

The prospective design is usually regarded as conceptually and methodologically more rigorous and accurate than the cross-sectional and retrospective designs. In principle, the prospective design provides a complete description of experience subsequent to exposure, including rates of progression, staging of disease, and natural history (Schlesselman, 1982, p. 18). Second, the design allows for the study of multiple potential effects of a given exposure, thereby obtaining information on benefits as well as risks associated with specified characteristics. Third, the procedure permits flexibility in choosing variables to be systematically recorded. Previously unsuspected variables displaying a strong association with the disease under investigation are more likely to be detected in the course of the study than if the focus is on a pre-defined set of variables. Fourth, the cohort method permits thorough quality control in the measurement of study variables (Schlesselman, 1982). It does so in part by reducing the possibility of bias in obtaining information from cohort members.

There are certain disadvantages to using the cohort method, however. Compared to the case-control design, the cohort study takes longer to conduct with a potentially long duration of follow-up. Because of the length of time involved, the possibility of confounding factors emerge as current practices or exposures to the study variables may change, making the research results irrelevant. One can reduce the length of the follow-up period by increasing the size of the cohort, but this alternative may involve considerable additional expense. Even with a longer follow-up period and smaller cohort, prospective

studies involve more time, money and personnel than do cross-sectional and retrospective studies. The potential for improved quality control also increases the potential expense of the project. For studies of rare diseases, a cohort design requires the use of a large number of subjects, also adding to the increased expense. Finally, maintaining follow-up on the cohort can be difficult as many subjects change residence or drop-out entirely from a study.

DATA FOR PSYCHOSOCIAL STUDIES

Rates and Scales

As noted above, psychosocial epidemiology is characterized by its use of psychometric data in conjunction with the clinical data typically employed in other epidemiologic studies. These two sets of data are distinguished by their source and forms of measurement. Clinical data usually comprise a nominal set of episodes of disease morbidity or mortality. They are usually obtained from hospital records, physical examinations, or death registries but may also be obtained through the administration of various screening scales in which an arbitrary cut-off point is used to distinguish between cases and non-cases or controls. These episodes are commonly expressed in terms of rates.

A rate is a statement of probability of the phenomenon measured per unit of population per unit of time. In epidemiology, this statement of probability involves three different items of information: (1) the number of persons experiencing the episode of disease morbidity or mortality, expressed as the numerator; (2) the population at risk from which the sample of cases are drawn, expressed as the denominator, and (3) a specification of the time interval. Thus, a mortality rate attributed to alcoholism of 41.4 per 100,000 Alaskan Natives in 1970 reflects a probability of alcohol-related deaths in a specific population for a designated year and is based on the number of Alaskan Native deaths from alcoholism in 1970 divided by the Alaskan Native population in that year. While the actual Alaskan Native population in 1970 was much smaller than 100,000, the use of this number (or a smaller factor of ten) serves to standardize the rate, allowing comparisons across different groups to be made.

The psychometric data utilized in psychosocial epidemiologic studies comprise a set of information on both dependent and independent variables. Levels of stress, for example, are often represented in psychometric data. These data are usually obtained through one or more scales measuring life events, psychological distress, mental disorders, psychophysiological symptoms, or personality characteristics. Several scales have been developed to discriminate between patients with varied psychiatric diagnoses and nonpatients for use in epidemiologic research in the general population. These scales are very similar in content and tend to correlate with each other. However, they usually provide only a general measure of psychological distress which is analogous to measures of temperature in physical medicine (Dohrenwend, Shrout, Egri, and Mendelsohn, 1980, p.1229). Other, more specific, scales are applied to diagnose specific forms of psychopathology.

One of the earliest and most widely used objective measures in psychosocial studies was a

22-item screening instrument developed in connection with the Midtown Manhattan study to provide an approximation of the degree of psychiatric impairment in community surveys. (Langner, 1962). The scale does not classify individuals according to type of psychiatric disorder but provides a rough indication of where people lie on a continuum of impairment in functioning due to common types of psychiatric symptoms. A similar although less widely used measure is the Health Opinion Survey used in the Stirling County study (MacMillan, 1957). Both measures have as their core a portion of the items from the Psychosomatic Scale of the Neuropsychiatric Screening Adjunct (NSA) devised as an aid to Selective Service screening during World War II (Dohrenwend & Dohrenwend, 1982, p. 1272) and the Minnesota Multiphasic Personality Inventory.

Dependent Variables

The major dependent variables in psychosocial epidemiology are illness and stress, although stress can also be viewed as an independent or intervening variable as well. Illness episodes can be expressed in the forms of rates or in terms of psychiatric scales while levels of stress are usually measured in terms of life events, patterns of behavior, psychophysiological symptoms, or biochemical indicators. The examination of medical data relating to the pattern and distribution of physical and mental illness is a critical part of psychosocial epidemiology. Not all studies examine the same types of data, however, and there is much overlap between the data examined in some studies which may be referred to as psychosocial and data examined in other types of epidemiologic investigations. This section will briefly examine three types of dependent variables. mortality, morbidity and stress.

Mortality Although mortality statistics are usually examined in other types of epidemiologic studies, they have played an important role in psychosocial studies as well. Several studies, for instance, have examined the effect of different forms of social support on health status by selecting a cohort, distinguishing members of the cohort on the basis of specified psychosocial variables, and then following the cohort over a period of time (e.g., Berkman and Syme, 1979; Wingard, 1982; House, Robbins and Metzner, 1982). Death certificates obtained from a state death registry provide data on the number and causes of death for cohort members throughout the time period. Death certificates and registries can also be employed for studies of potential social factors affecting survival rates for particular diseases such as cancer (Dayal and Chiu, 1982) or coronary heart disease (Chandra, Szklo, Goldberg, and Tonascia, 1983).

Obtaining the data on mortality is only the first step in their use in an epidemiologic study. To quantify the data, the number of deaths in a population is expressed in terms of a rate. A mortality or death rate requires the following information: (1) the number of deaths in the exposed or affected population during a certain time period, which comprises the numerator; (2) the total population exposed to the risk of death, comprising the denominator; and (3) a specified period of time, usually a 1-year interval. Expressed formally, the annual death rate is as follows:

$$\text{Annual death rate (ADR) from all causes (per 1,000 population)} = \frac{\text{Total number of deaths occuring in a one year period}}{\text{Number of persons in the population at mid year}} \times 1{,}000$$

The units of time and population may be selected by the investigator to suit his own purpose, but they must be specified. Death rates can also be made specific for a variety of characteristics, such as age, cause of death, marital status, race and occupation. Age-specific rates are perhaps the most common form of specific rate, given the high association between age and mortality. Thus, a suicide rate of 19.2 per 100,000 among 18-24 year olds in the United States in 1977 would be calculated as follows:

$$\text{ADR from suicide for persons aged 18-24 (per 100,000 population)} = \frac{\text{Number of suicides of persons 18-24 during 1977}}{\text{Number of 18-24 year olds in the United States as of July 1, 1977}} \times 100{,}000$$

Another type of rate frequently used in the "case fatality rate":

$$\text{Case Fatality Rate (\%)} = \frac{\text{Number of individuals dying during a specified period of time after onset or diagnosis of disease}}{\text{Number of individuals with the specified disease}} \times 100$$

This rate represents the risk of dying during a definite period of time for individuals diagnosed as having a particular disease. As with the death rate, the period of time during which the deaths occur should be indicated. Case fatality rates can also be made specific for age, sex, severity of disease, or any other factors of clinical and epidemiologic importance.

Morbidity More commonly utilized in psychosocial epidemiologic studies are morbidity statistics for both physical and mental disorders. Because of its association with stress and the relative ease in obtaining the data by investigators, blood pressure levels are commonly utilized in psychosocial studies (Harburg et al., 1978; Reed, McGee, Cohen, Yano, Syme, and Feinleib, 1982; Keil, Tyroler, Sandifer, and Boyle, 1977; Hypertension Detection and Follow-up Program Cooperative Group 1977). Usually, three readings are taken at any one point in time by someone trained in the standard blood pressure technique and a mean average used as the reading for each subject. Blood pressure readings can be taken independently or in conjunction with a physical examination.

Physical examinations are another source of data for psychosocial investigations. Several studies have relied upon regularly scheduled medical exams by physicians in studies of coronary heart disease (Marmot and Syme, 1976; Reed et al., 1982; Haynes and Feinleib, 1980); ulcers (Cobb and Rose, 1973), cancer (Fox, 1978) and so on. These examinations may be comprehensive, including a complete biochemistry and EEG workup, or they may be specific to the disease itself, the former method being preferred because of the potential for numerous confounding factors.

While many if not most studies involve physical examinations by study investigators, morbidity data can also be obtained from hospital records. Dayal and Chiu (1982), for

instance, employed a tumor registry from a university hospital in their study of racial differences in survival for prostatic cancer.

Interview surveys are also employed to obtain morbidity data. These questionnaires usually contain details of past medical history, smoking history, physical activity, and dietary habits. Several of these interview surveys such as the Health Interview Survey (HIS) and the National Ambulatory Medical Care Survey (NAMCS) have been employed in nationwide studies; others have been used on smaller samples. An example of such a survey is the Seriousness of Illness Survey (Wyler, Masuda and Holmes, 1968), a self-report checklist of 126 commonly recognized physical and mental symptoms and diseases. This instrument includes a general severity weight for each disorder, reflecting prognosis, duration, threat to life, degree of disability, and degree of discomfort. This instrument has served as a useful tool in several stress and illness studies (e.g., (Kobasa, Maddi and Courington, 1981; Dohrenwend and Dohrenwend, 1974). Other survey instruments are specific to particular disorders, such as the London School of Hygiene Cardiovascular Questionaire used by Marmot and Syme (1976) in their study of coronary heart disease and acculturation among Japanese-Americans.

As with physical disorders, much of the epidemiologic research on psychiatric illness has relied on data obtained from screening scales, from records of hospital admissions, or from clinical evaluations by study investigators. Screening scales are perhaps most frequently used in such studies. Several protocols have been developed to assess psychopathology in general such as the Minnesota Multiphasic Personality Inventory, the Hopkins Symptom Checklist, and the Present State Examination (PSE), while others are designed for specific forms of psychopathology such as schizophrenia (SADS-RDC scale), or depression (CES-D scale).

Another source of data on psychiatric illnesses has been derived from hospital records. In his review of studies of mental disorder and social class, Fried (1969) notes that records of psychiatric hospitalization are most often used as data and rates are usually based on the number of patients in a demographic category as the numerator with census calculations of the population at risk in these categories as denominators.

Clinical assessments are also used frequently in assessments of psychiatric disorder. Vaillant (1976), for example, using data collected in a follow-up study of college males, went through autobiographical responses to biennial questionaires, interviews, summaries, psychological tests and protocols to record each man's behavior in a time of crisis in the form of a vignette. Vignettes were then grouped into clusters on the basis of the type of defense mechanism employed by each study subject. The assessments of defense mechanisms were then verified in a blinded evaluation by two other judges. The judges independently rated each of 50 cases, labelling vignettes according to a glossary of 18 defenses.

As with instances of mortality, the quantification of events of illness or disease takes the form of rates. There are two major types of morbidity rate: incidence and prevalence. The incidence rate of a disease is the number of new cases of the disease occurring within

a specified population during a specified time period.

$$\text{Incidence rate per 1,000} = \frac{\text{Number of new cases of a disease occuring in a population during a specified period of time}}{\text{Number of persons exposed to risk of developing the disease during that period of time}} \times 1{,}000$$

In order to establish the true incidence of any disease three basic conditions have to be met (Lemkau and Crocetti, 1958). First, the identification of the disease or illness under study should be objective and reliable. Second, all cases should be accounted for, either by recording every case which occurs in the study population or by knowing the ratio of unknown to known cases. Third, the population from which the cases are drawn should be clearly defined and carefully enumerated. While many studies are content to rely upon decenniel census statistics in calculating the population at risk, more accurate means of enumeration are available and should be used whenever possible.

In studies where a cohort is being followed and observed over a period of time, it is often more useful to employ person-years of observation as denominators in the computation of rates. This is particularly true in studies where subjects enter and leave for varying periods of time. Moreover, the age distribution of the groups under observation will change over the course of the follow-up period, as will morbidity and mortality rates. The use of person-years makes it possible to express in one figure the period when a varying number of persons is exposed to the risk of a disease episode. Incidence rates using person-years as the denominator would take the form represented in Table 1.2. Both the number of individuals in the cohort and the duration of observation of each person are taken into account. "For example, five persons who remain under observation for twenty years contribute one hundred person-years, as would one hundred persons observed for one year" (Lilienfeld and Lilienfeld, 1980, p. 245).

Table 1.2

Calculation of Person-Year Incidence for Disease Events by Characteristic or Exposure in a Follow-up Study

Characteristic or Exposure Groups	Number of Individuals	Person Years	Number of Events	Person-Year Incidence Rate*
A	500	1,600	5	3.1
B	500	2,400	7	2.9
TOTAL	1,000	4,000	12	3.0

* per 1,000

Prevalence rates measure the number of cases present at, or during, a specified period of time. The prevalence rate equals the incidence rate times the average duration of the disease. For example, if the average duration of hypertension is three years and its incidence rate is 15 per 1,000, the prevalence rate would be 45 per 1,000.

$$\text{Prevalence rate per 1,000} = \frac{\text{Number of cases of disease present in the population at a specified time}}{\text{Number of persons in the population at that specified time}} \times 1{,}000$$

The two types of prevalence rates used in epidemiologic studies are point prevalence and period prevalence. Point prevalence refers to the number of cases present at a specified moment in time, such as the date the sample is drawn and the information obtained; period prevalence refers to the number of cases that occur during a specified period of time, for example, a year. Period prevalence consists of the point prevalence at the beginning of the specified time interval plus all new cases that occur during the interval.

The use of and differences between prevalence and incidence rates can be demonstrated using data obtained in the Honolulu Heart Program study of coronary heart disease and acculturation among men of Japanese descent (Reed et al., 1982). A sample of 4,653 men of Japanese ancestry were observed over a seven-year period from August 1, 1971 to January 1, 1979. The total number of cases of coronary heart disease (defined as a clinically diagnosed episode of myocardial infarction or angina) observed throughout the period was 482. Of this figure, 264 cases were recorded at the beginning of the study, August 1, 1971. The point prevalence of coronary heart disease among the subjects, therefore, was

$$\text{Point prevalence (per 1,000 as of August 1, 1971)} = \frac{264}{4653} \times 1{,}000 = 57$$

The incidence rate was calculated by taking all subjects free of the disease at the beginning of the study (4653 - 264 = 4389) and observing 218 new cases of coronary heart disease among this group throughout the seven-year period. The incidence of coronary heart disease among this sample of Japanese men, therefore, was

$$\text{Incidence rate (per 1,000 1971-1977)} = \frac{218}{4389} \times 1{,}000 = 50$$

The period prevalence rate would be calculated by taking the total number of cases of coronary heart disease observed throughout the entire period and the total population at risk. The prevalence of coronary heart disease for the seven-year period, therefore, would be

$$\text{Period Prevalence (per 1,000 1971-1977} = \frac{482}{4653} \times 1{,}000 = 103.6$$

Both incidence and prevalence rates can be made specific for age, sex and/or any other social or psychological characteristics.

Of the two types of rates, incidence rates are preferred in psychosocial epidemiology because they provide a true index of the risk of disease for specific populations. However, both are faced with certain types of disadvantages. Incidence rates, for example, are affected by numbers, types, and availability of services. As will be discussed in greater detail below, the variability in these services for different segments of a population can bias the statistical analyses. On the other hand, prevalence rates tend to

overweight problems of long-term duration and underweight those of short-term duration.

Stress Because stress is defined in terms of non-specific tendencies, it cannot be measured directly but rather must be assessed in terms of human response. Baum, Grunberg and Singer (1982) have grouped the measures of human response used in psychosocial studies into four broad categories: self-report, performance, psychophysiological and biochemical. Self-report measures involve the use of questionaires and scales to obtain information on experiences, perceptions, and evaluations. Performance measures record gross behavioral changes in response to activities believed to be stressful. Psychophysiological measures assess organ or system function in the body, such as cardiovascular or electrodermal response, muscle tension, and respiration. Biochemical measures of stress (e.g. catecholamine and corticosteroid secretion) examine levels of activity in the endocrine system (Baum et al 1982:219).

Several scales exist which index the experience of stress based on data obtained from interviews or questionnaires. Many of the scales which screen for psychological impairment have already been discussed. These scales usually ask subjects to list psychological or somatic experiences of distress and to evaluate their intensity. In a study by Turner (1981), for instance, psychological well-being was measured through use of the Brief Symptom Inventory (BSI) scale, composed of 53 items, each item being rated on a five point distress scale. The BSI is scored in terms of nine primary symptom dimensions: somatization, obsessive-compulsive, interpersonal sensitivity, depression, anxiety, hostility, phobic anxiety, paranoid ideation, and psychoticism. Other self-report scales of stress include the General Health Questionnaire used by Andrews, Tennant, Hewson, and Vaillant (1978), and a perceived strain index developed by Pearlin and Schooler (1978), consisting of a series of Lickert-type items encompassing financial, marital and work-related sources of strain.

One of the most common measures of stress, however, is the number and types of life events preceding an illness event. Life events are treated here, however, as independent or intervening variables for three reasons. First, life events are presumed to occur prior to a stressful response although the causal relationship may occur in the opposite direction as well. Second, the same life events do not necessarily produce the same response in all individuals. Third, different individuals may give similar responses to entirely different events. More will be said about the relationship between life events and stress below.

Performance measures of stress rely on such indices as loss of coordination, fatigue (Appley and Trumble, 1967), increased reaction time (Frankenhaeuser, 1975), and problem solving (Glass and Singer, 1972). These measures involve exposure to a stressful task or situation and comparison of response between exposed and non-exposed subjects. A number of tasks have been identified for this purpose and measures have been developed to assess the stress experience. One such measure has been provided by Glass and Singer's (1972) research on noise and urban stress. Subjects exposed to stressful conditions generally

perform more poorly on proof-reading tasks requiring great concentration than do subjects exposed to less aversive conditions (Glass and Singer, 1972; Cohen, 1980). This differential performance seems to reflect a cost of coping with the stressful condition (Baum et al., 1982, p.223).

Psychophysiological measures of stress are concerned with degrees of arousal or activation associated with the sympathetic nervous system (SNS) (Baum et al., 1982:223). Blood pressure, respiration rate, heart rate, muscle tension, and galvanic skin response have been used in numerous studies as indices of stress.

Two major biochemical indicators serve as reliable measures of stress in psychosocial studies,urinary and plasma adrenal corticalsteroid levels and urinary and plasma catecholamine levels. Each represents the response of a different endocrine system which react to different perceptions of the body and different environmental factors. Several excellent reviews of research on the hypothalamic-pituitary-adrenal cortex axis and sympathetic-adrenal medullary axis (Mason, 1968a, 1968b, 1975; Brown, 1981; Henry, 1982; Baum et al., 1982) which document extensively the evidence of psychosocial influences on activity in these systems. Corticosteroids are secreted by the adrenal cortex, and serve many important functions including regulation of body electrolytes and water, carbohydrate metabolism, protein and fat metabolism, and ability to tolerate stressors. The 17-hydroxycorticosteroids (17-OCHS), primarily cortisol, have been repeatedly demonstrated to be reliable indicators of stress. Numerous studies have found increased levels of 17-OCHS in urine and blood samples of subjects under stressful situations (Mason, 1975; Baum et al., 1982).

Catecholamines such as epinephrine and norepinephrine are released by the adrenal medulla under situations of stress and produce numerous physiological reactions, including increased cardiac output and blood pressure, increased minute volume of respiration, glycogenolysis, and mobilization of free fatty acids, stimulation of the pituitary to release ACTH, diversion of blood flow from the digestive system to the skeletal muscles, and central nervous system effects such as desensitization to pain and increased alertness (Brown, 1981, p. 77). Increased levels of catecholamines in urine or blood samples serve as reliable indicators of stress and have also been used in numerous studies.

Independent Variables

Sociodemographic characteristics Whether one is attempting to understand the etiology of stress or particular stress-related disorders, one or more independent variables must be taken into consideration and controlled for. Demographic characteristics such as age, sex, residence, race, ethnicity, marital status, and socioeconomic status (SES) have been the subject of several different psychosocial epidemiologic studies. Most of these variables are straightforward and are relatively easy to operationalize. Others present greater difficulties. Socioeconomic status, for instance, has been operationalized in several different ways, making comparison between studies difficult. Education, income, occupation, and census tract of residence have all been used, independently or in combina-

tions, as indices of socioeconomic status. Several scales of SES exist including the Duncan Socioeconomic Index, the Green Manual, and the Hollingshead two-factor Index of Social Position. Usually a two factor-index of social position, measured by education and occupation, is employed in epidemiologic studies. As McQueen and Siegrist (1982, p. 353) observe, this index is used "because it is a 'quick and dirty' approximation of class which appears to satisfy many epidemiologists and is regarded as easy to interpret."

In studies of race, skin color has been used to examine within-group differences in health (Harburg et al., 1978; Keil et al., 1977). Ethnic group is usually obtained from self-reports of study subjects. Other indices of ethnic group status used in epidemiologic investigations have been race, surname, language spoken, and birthplace of self, parents and grandparents (Hayes-Bautista, 1983).

Life Events As noted above, life events are usually regarded as indicators of the degree of stress experienced by a particular individual or groups of individuals. Thus, although the stress itself can be said to be a response to one or more events, the events themselves are used as an index of the degree of stress. Among the earliest and most widely used measures of stress based on the concept of life events are the Schedule of Recent Experiences (SRE) and the Social Readjustment Rating Scale (SRRS) developed by Holmes and Rahe (1967). The SRE is a list of 43 life events found to involve some sort of adjustment for most people. The scale measures the intensity and length of time necessary to accommodate to a life event, regardless of the desirability of the event. The SRRS gives different weights to each life event. Subjects are asked to rate a series of predefined events on the basis of personal experience. Marriage is given an arbitrary value of 500 and other events are scaled relative to this life event. The amount of life change is derived from the multiplication of the weight for each event by the frequency of its occurrence for the subject during the time period under study. This product is expressed as Life Change Units (LCU). Pearson product-moment correlations were used by Holmes and Rahe to determine that a universal agreement between groups and among individuals about the significance of life events under study transcends differences in age, sex, marital status, education, social class, generation American, religion and race. Several epidemiologic studies have been conducted employing the scales developed by Holmes and Rahe.

In addition to the Schedule of Recent Experience and the Social Readjustment Rating Scale, otner scales have been devised to measure the effect of life events on psychological distress and illness. In a study of Antonovsky and Kats (1967) a measure of life events was obtained by eliciting responses to 30 questions, classified into one of four subject areas: physical trauma (e.g., illness, operation, beating); change in general environment (e.g., move from village to town, migration, internment in camp); changes affecting primary interpersonal relations (e.g., marriage, death of parents); and changes in status (e.g., type of work, employment). Within each of these areas, five levels of crisis intensity were distinguished. Patients and controls were compared on the basis of six predefined scores:

1. A total crisis score (TCS) representing the number of crisis points of an individual, irrespective of crisis area, age, or number of crisis.

2. A sub-TCS within each of the four crisis areas.
3. A serious crisis score representing the number of crises scored as four or five (NSC).
4. A sub-NSC within each of the four crisis areas.
5. A crisis score for two, five and ten years before the "age at onset" and from birth till age 15.
6. A "crisis concentration" score, defined as 15 crisis points within a five-year period.

Using this procedure, Antonovsky and Kats found that patients differed significantly from controls with respect to total crisis score, serious crisis scores, total and serious crisis scores in the areas of environmental change and interpersonal relations, and cumulative crisis scores in two, five and ten years before onset.

While scales such as these measure life events strictly in terms of the number of events, other scales measuring the quality of events have been developed. Dohrenwend (1973), for instance, developed an index of the undesirability of events experienced to test the hypothesis that the type of life event was more salient in the experience of distress than the number of events. In this index, all events reported by the respondents were scored as culturally defined losses or gains according to the following schema: a loss was defined as an event or change generally considered to be undesirable and scored as +1; a gain was defined as an event or change generally considered to be desirable and scored as -1; and ambiguous events or changes were those whose desirability was subject to debate, scored as 0. The events were coded by two coders working independently and each individual's score was the algebraic sum of all the events he reported.

Related to this issue of relative importance of different life events, efforts have been made to weight life events. As noted above, Holmes and Rahe (1967) used life change units (LCU) to weight the amount of adjustment required by each life change. This measure was obtained by asking a standardization sample how much disruption in behavioral patterns would result from each event. An alternate procedure has been to have individuals indicate the level of idiosyncratic disruption the event caused them. Some researchers (e.g., Byrne and Whyte, 1980; Sarason, Johnson and Siegel, 1978) have argued convincingly for ideosyncratic weighting, since an event may have a different impact on one person relative to another (Newcomb, Huba and Bentler, 1981, p. 401). However, others (Dohrenwend, 1980; Kessler, 1979) have argued that weighting confounds relationships between life events and illness.

One index which incorporates the notion of a life event and social status is level of acculturation. In their study of coronary heart disease among Japanese and Japanese-Americans, Marmot and Syme (1976) developed an Ethnic Identity Scale, a questionaire which included three indices of acculturation: (1) culture of upbringing, (2) cultural assimilation and (3) social assimilation. Culture of upbringing combines information on years spent in Japan, age left parents' home, ever lived on a farm, where schooling took place (Japan or U.S.), years spent in Japanese language school, religion while growing up, friends while growing up, and wife's cultural background (if married). Cultural assimilation combines information on ability to read Japanese, frequency of speaking Japanese to wife, children, and friends. Social assimilation is indicated by ethnicity of profession

and ethnicity of co-workers.

Alienation is another phenomenon which incorporates elements of life events and status. In a study conducted by Seeman and Anderson (1983), the concept of alienation was operationalized using a series of different scales which measured such items as powerlessness, degree of social integration, job satisfaction, occupational striving, and career mobility. Powerlessness is often measured in psychosocial studies by variations of the Locus of Control Scale developed by Rotter (1962). These scales consist of a series of questions designed to elicit a measure of operationally defined concepts. They represent questions from larger scales which are subjected to a factor analysis in order to identify questions associated with similar concepts. These questions then form a smaller scale designed specifically to measure the construct desired. Factor loadings also indicate the extent to which a question is salient to the construct being measured.

Social Support One of the most heavily studied issues in the field of psychosocial epidemiology has been the impact of social supports on health and well-being. Berkman and Syme (1979) examined four sources of social contact: (1) marriage, (2) contacts with close friends and relatives, (3) church membership, and (4) formal and informal group associations. From this information, a Social Network Index was constructed, scoring subjects on a scale from I (fewest social connections) to IV (most connections). In a similar study conducted by House et al. (1982), the measures of social relationships and activities were obtained from self-administered questions in four major categories: (1) intimate social relationships (marital status, visits with friends and relatives, going on pleasure drives and picnics); (2) formal organizational involvements outside of work (going to church or meetings of voluntary associations); (3) active and relatively social leisure (going to classes or lectures, movies, plays, fairs, museums, etc.); and (4) passive and relatively solitary leisure (watching television, listening to the radio, reading).

Other methods of obtaining data on social support networks have been employed in psychosocial epidemiologic studies. Turner (1981) employed a story identification technique developed by Kaplan (1977) composed of seven vignettes. The Sense of Support Scale developed by Panagis and Adler (c.f., Aneschensel and Stone, 1982) consists of a series of Lickert-type items encompassing two types of support: socioemotional support (e.g. thoughtfulness, understanding) and instrumental help (e.g., assistance with work or with problems). Respondents are asked how often during a specified period of time someone has provided them with the listed types of support. Response categories range from 1 ("not at all") to 4 ("very often").

Personality Finally, measures have been developed which assess the coping style of an individual to determine its effect on the experience of psychological distress. One such measure, developed by Vaillant (1976) and used in other studies (Andrews et al., 1978) consists of questions designed to measure the maturity of the habitually used coping style. Respondents are presented with two situations with six responses listed for each situation, each situation to be rated as "like" or "not like" me. A coping style score is calculated by giving a score of +1 for each "like me" response on the items characterizing

mature ego defense mechanisms and a score of -1 for each "like me" response on the items characterizing immature defenses and then summing these scores (Andrews et al., 1978, p. 309).

Other measures have been developed which examine specific components of ego strength. Pearlin and Schooler (1978), for example, examined three personality characteristics used by individuals to help them cope with threats posed by events and objects in their environment: self-esteem, self-denigration, and mastery. Data on these variables were obtained from a series of questions on an interview schedule administered to subjects. Responses to the questions were factor analyzed and scored to provide a measure of psychological coping resources. Kobasa et al. (1981) employed six different scales to measure the degree of hardiness in subjects: alienation from self and alienation from work scales of the Alienation Test (Maddi, Kobasa and Hoover, 1979); a security scale of the California Life Goals Evaluation Schedule (Hahn, 1966); a cognitive structure scale of the Personality Research Form (Jackson, 1974); the external locus of control scale (Rotter et al., 1962) and the powerlessness scale of the Alienation Test (Maddi et al., 1979).

Limitations of Data

Among the major tasks involved in a psychosocial epidemiologic study are the recognition and handling of the limitations of the data. Validity and reliability of analyses can be compromised by the nature of the disease or stress episodes themselves, the methods used to obtain the data, the spurious relationships resulting from false or defective data, and the indirect associations resulting from undetected confounding variables. This section will examine these problems and suggest methods for handling them. Methods for handling possible spurious relationships and indirect associations also will be discussed in the section on data analysis.

Problems of Definition Many of the data limitations in psychosocial studies are associated with the nature and definition of mental illness, the sources of data, the measures of stress and life events, and the operationalization of independent variables. Perhaps the most basic problem with the data on disease episodes used in psychosocial studies is the lack of a precise definition of the disorder itself. This is a common problem in psychiatric epidemiology where changes in diagnostic policy and nomenclature can result in alteration of rates for various mental disorders. Dohrenwend and Dohrenwend (1982), for example, point to changes in nomenclature as being partially responsible for the increase in the rates of psychiatric disorders since World War II. Cross-cultural epidemiologic studies are especially subject to this limitation because the episodes can vary in content even though the form of psychiatric disorders is essentially constant throughout the world (Kiev, 1972).

This problem is not limited to psychiatric disorders, however. Comparisons of studies of hypertension, for example, are often hindered by the fact that different criteria are employed in defining a "case" of hypertension. As McQueen and Celentano (1982) note, some

studies have used blood pressure levels of 140/90mm Hg as their cutting point for defining hypertension while other studies have used levels as high as 160/95mm Hg. Alcoholism is similarly beset with problems of definition. Differences can exist when frequency of drinking (number of drinks per unit of time) is used as a criterion for case definition and when amount consumed (measured in terms of amount purchased, blood alcohol levels, and so on) is used. The inconsistencies of these definitions lead to underenumeration or overenumeration of the disease episodes.

Inconsistency of case definition is due to other factors as well. Different diagnoses for similar phenomena may result because of the subjective nature of the diagnostic process for some disorders. Adebimpe (1981), for instance, notes four possible sources of misdiagnosis in clinical examinations of Black patients: the social and cultural distance between patient and clinician, stereotypes of Black psychopathology, false-positive symptoms, and biased diagnostic instruments. A similar problem of misdiagnosis emerges with respect to social class. Mischler and Scotch (1963) point to evidence suggesting that lower-class patients are often given diagnoses for more severe illnesses than their middle-class counterparts, even though the symptoms are similar. Hospitals and clinics catering to middle and upper class patients may also underdiagnose particular disorders such as schizophrenia because of the social stigma involved. Even when patient and clinician come from similar cultural or social backgrounds, the subjective nature of data collected by clinical impression can be a constant source of error.

To minimize the problems associated with episode definition, therefore, two separate perspectives are required. The first perspective is that of the investigator. Agreement among investigators as to what constitutes a "case" of a particular disease should be sought. The definition of a case should have two specifications (Schlesselman, 1982): the establishment of objective criteria for the diagnosis of the study disease and a statement of eligibility criteria for the selection of individuals to participate in the study. The second perspective is that of the subjects. Objective and subjective criteria as to case definition may differ because of differences in perception and evaluation of symptoms (Kleinman, Eisenberg and Good, 1978). Hence, explanations for symptom occurrence and an account of its subjective significance should be obtained from study subjects.

It is suggested that, whenever possible, multi-method strategies of case identification and diagnosis be employed in psychosocial studies. Dohrenwend and Dohrenwend (1982) outline how such a procedure would work:

> For example, a self-report interview like the Psychiatric Epidemiology Research Interview (PERI), based on a psychometric approach to measuring dimensions of psychopathology, could be used economically to screen samples from the general population. Such screening would yield subsamples of individuals with various types of severe symptomatology. Individuals screened by high scores on the screening scales could then be followed up in a second stage of the research and interviewed by experienced clinicians with diagnostic instruments like the Schedule for Affective Disorder and Schizophrenia (SADS) or the Present State Examination (PSE) to provide rates for particular types of disorders. Such a two-stage procedure could capitalize on the ability of a psychometric instrument to provide reliable measurement over the full range of important dimensions of symptoms of psychopathology and on the abilities of a clinical examination to provide reliable diagnoses of individuals in groups where the types of symptomatology involved are not rare (pp. 1275-1276).

Moreover, these combinations of screening scales and clinical exams help establish

validity of data.

Source of Data Another set of limitations of data on disease episodes is associated with the source of information. The number of cases of disease episodes are usually taken from records of hospital admissions and diagnostic instruments. The use of these instruments in general epidemiologic studies is hindered by the fact that they have been developed on psychiatric patients or specific groups which cannot be regarded as typical of the general population. The methodological "net" used to detect cases of psychiatric disorder, therefore, may be too fine, detecting only a certain type of the case under investigation.

Second, the use of records of hospital admissions will bias the number of cases in several respects. With respect to psychiatric disorders in particular, individuals recorded as hospital admissions may not be representative of those who have experienced the disease episode under investigation. For example, Mischler and Scotch (1963) note that hospital admissions for schizophrenia are representative only of those cases which do not recover quickly. Use of these admissions, therefore, would produce distorted rates of both incidence and prevalence. Hospital admissions for other severe psychiatric disorders represent only those cases which are treated by clinicians. According to Fried (1969), ambiguities are built into the total situation in which the individual comes to psychiatric notice and is diagnosed. A common assumption is that severe disorders will come to the attention of clinicians and will inevitably be hospitalized. Unfortunately, this view is based on several unproven assumptions, including the views that psychoses are persistent and that community tolerance operates only within a narrow range, forcing the patient to seek treatment within the modern health care system.

The problem of using only treated cases in calculating rates of disease episodes pertains for all types of disease. An estimated 70% to 90% of all self-recognized episodes of sickness in general are believed to be managed outside the formal health care sector, usually employing popular or "folk" health care practices (Kleinman, Eisenberg and Good, 1978). These disease episodes go unrecorded, resulting in rates which are underenumerated.

Whatever the nature of the disease episode, however, comparison among exposed and nonexposed groups or cases and controls is difficult because of differences in numbers, types, and availability of services. Knowledge of, preference for, and access to health care and medical facilities are known to vary by age, sex, ethnic group, and socioeconomic status. Individuals with certain disorders may prefer not to seek professional medical care while others are restricted in the care available by virtue of lack of sufficient financial resources or distance from medical facilities. Thus, comparison of morbidity rates among different social groups may bias findings, particularly when the data is derived from a narrow range of sources such as state hospitals. Mischler and Scotch (1963) note that the use of state hospital admissions as the sole source of data for computing morbidity rates in different social groups might bias findings because other treatment resources such as private hospitals, clinics, and outpatient facilities are used dispro-

portionately by different social groups. Morbidity rates based on state hospital admissions would reflect only those patients unable to use other health care resources and hence be incomplete, overrepresenting certain social groups and underrepresenting others. This is particularly true for groups distinguished on the basis of socioeconomic status. Individuals in the higher class groups are relatively more likely to have had their illness detected early, and if hospitalized to have been released after a short period of treatment. A higher proportion of their admissions would then be readmissions. This factor would also tend to reduce the apparent incidence in higher class groups since incidence rates are based on first admissions.

Because assessment of health risks are often obtained by using hospital admissions, a spurious association could be obtained between various diseases and such psychosocial factors as psychological coping resources, socioeconomic status, race, and ethnicity if the groups under study do not share the same probability of hospital admission with or without the disease (Berkson, 1946). This spurious association, also known as "Berksonian bias," has been demonstrated to exist for numerous characteristics such as those listed above. Berksonian bias can indicate an association between a characteristic and morbidity where none exists, or it can have the reverse effect in that differences in hospital admission rates may conceal an association between the characteristic and a specific disease in a study,.even though one actually exists.

Finally, the level of record-keeping efficiency varies from hospital to hospital, producing statistics which do not accurately reflect the prevalence or incidence of disease in a particular hospital or in a particular community.

Bias in the counting of disease episodes can come from the researcher as well as the source. Non-clinical observations of disease episodes in the general population are subject to potential bias resulting from knowledge of the hypothesis or hypotheses being tested. According to McQueen and Seigrist (1982), this is particularly true in cases which are considered to be even slightly deviant, as in studies of alcohol use. An investigator eager to find a relationship between alcohol abuse and suicide, for example, may be biased in his perception of problem drinking and find instances of alcohol abuse more frequently among cases than among controls, even though the degree of drinking is similar in both groups. Such bias can be reduced, however, by utilizing two or more individuals to evaluate characteristics such as problem drinking, personality variables or measures of psychological distress. Quantitative techniques such as a Pearson product-moment correlation can then be employed to assess the degree of reliability in the judgements of the investigators.

Measures of Stress As described above, there are four general types of measures used in studies of stress: self-report, performance, psychophysiological and biochemical indicators. Each of these types possess a particular range of limitations. The type of measurement most frequently criticised in the literature is the self-report. The most common problem associated with self-report measures of stress is that of response bias. Some populations will say yes or no to questions regardless of content (Kennedy, 1973;

Dohrenwend and Dohrenwend, 1969). "Certain types of bias seem particularly likely to occur when psychiatric symptoms form the subject of the interview because such symptoms are by definition subjective and because psychiatric conditions, particularly if they result in institutionalization, still carry a social stigma" (Vernon, Roberts and Lee, 1982, p. 483). This problem occurs, for example, when collecting data on alcohol intake where, as McQueen and Celentano (1982) note, survey respondents will generally underreport their consumption. In other instances, responses tend to be overdetermined where a positive response to specific questions in the measurement of personality in general and psychopathology in particular can be given by different individuals for different reasons, some of which have less to do with the presence of the characteristic being measured than such extraneous factors as differences in how the individuals understand the meaning of such words as "depressed" or "anxious" (Dohrenwend and Dohrenwend, 1982). As a result, self-report scales may have the effect of measuring different things for different groups. The Center for Epidemiologic Studies - Depression (CES-D) scale, for instance, has been shown to produce different responses in men and women, leading to the conclusion that the scale may not be measuring the same phenomenon among the two groups (Clark, Aneschensel, Fredrichs et al., 1981). Response tendencies have been shown to vary systematically across factors such as age, gender, ethnicity, and socioeconomic status. If the objective of the scale is to compare subgroups in order to study etiologic factors, it must be constructed so as to measure this phenomenon similarly in all groups (Vernon, Roberts and Lee, 1982).

Response bias can result from several factors. One is the differential ability of individuals to evaluate their own symptoms. These measures tap conscious experience but it is the experience of illness and not disease and "illness is shaped by cultural factors governing perception, labeling, explanation, and valuation of the discomforting experience" (Kleinman, Eisenberg and Good, 1978, p. 252). Self-reports of seemingly obvious indications of changes in physiological or behavioral state will be subject to considerable individual and cultural variation.

Another factor contributing to response bias is the problem of selective recall. Monroe (1982) found that as much as 60% of life events recalled by subjects in a retrospective procedure may be underreported for even the most recent four-month retrospective period. Additionally, particular types of events (desirable events) may be relatively more susceptible to such reporting distortion.

A third factor contributing to response bias is the mediation of self-reports by the coping style of the individual. Use of particular psychological defense mechanisms in an effort to cope with stress such as denial or repression can result in distorted answers to questionnaires (Baum et al., 1982).

Two separate options are available to minimize the problem of overdetermined responses to interview questions. The first, advocated by Wing, Cooper and Sartorious (1974) is to employ cross-examination in the clinical interview. A second option, advocated by

Dohrenwend and Dohrenwend (1982), is to sum conceptually related items to form scale scores. Consistency among the responses to a series of questions would indicate the scale's reliability in measuring stress or illness despite the error in responses to any particular question. Mathematical procedures such as Cronbach's alpha or Spearman's r can be used to estimate the degree of internal consistency reliability achieved with such a scale. A scale may be viewed as "reliable" in a subsample if it has a value of Cronbach's alpha of .50 or better.

Performance measures are limited in three respects. First, there is the difficulty in differentiating between stressed and nonstressed subjects. Both groups may exhibit similar responses for entirely different reasons. Second, the stress observed is context-specific, limiting the range of conclusions that may be drawn. The artificial nature of the stressful context must be taken into account when attempting to generalize results to the wider population. Third, the gross nature of the indices leaves open questions as to the precise nature of the mechanisms involved. Stress-induced behavioral changes reflect coping activity (Lazarus, 1966) and are strongly influenced by the context in which stress occurs and the individual's previous experience, value system, personality, and knowledge (Brown, 1981).

Physiological measures are usually regarded as being highly reliable. They, depend, however, on the accuracy of the equipment and skill of the investigator in using the equipment. In certain situations, as in the event of a naturally occurring stressor involving trauma or severe loss, use of such instruments may be impractical, creating additional stress for the subject. Even in less severe circumstances, blood pressure, heart rate, and muscle tension can all be affected by factors other than general stress conditions (e.g., activity, altitude, diet, genetic factors, smoking) which need to be controlled for.

Biochemical measures are usually regarded as being the most reliable indicator of stress and the easiest to obtain. Yet, even these measures are limited in certain respects. Interpretation of corticosteroid and catecholamine levels is not always straightforward (Baum et al., 1982). For instance, in using 17-OCHS levels as a measure of adrenal cortical response to stress, one must always take into account large individual differences in coping styles and exposure to certain stressors (e.g., Mason, 1975).

A similar problem occurs with respect to the use of catecholamine levels as indices of stress. Increases in catecholamine levels are not necessarily related to stress but may be caused by the ingestion of coffee, tobacco, or alcohol (Henry, 1982). Catecholamine levels can vary depending on the type of sample drawn. Plasma samples, for example, are more problematic than urine samples because fluctuations in plasma catecholamine levels are rapid and extremely sensitive to movement. On the other hand, epinephrine and norepinephrine excretion in urine samples display a circadian pattern and hence levels will depend upon the time of day the sample is drawn. (Baum et al., 1982).

Thus, while biochemical markers of stress have certain advantages over self-report scales

and performance measures, they should be used in conjunction with psychological protocols. It is relatively easy to gatner the required samples for bioassay (e.g., urine) and, combined with self-report measures, would provide a better understanding of stress response.

Operationalizing Independent Variables The fourth set of limitations to the data in psychosocial studies concerns the problems involved in operationalizing independent and intervening variables. These problems are conceptual as well as methodological in nature. Life events are particularly representative of these problems. One concern in the measurement of life events and their use as predictors or indices of stress has been the issue of quality versus the quantity of stressful events. Life events have been the focus of considerable debate regarding the relative importance of the number of events versus the type of events. The Holmes-Rahe scale is based on the notion that change itself, whether positive or negatively valued, is the critical factor in the development of stress while other studies (e.g. Brown and Birley, 1968; Johnson and Sarason, 1978) have placed greater emphasis on the quality of events. The available evidence appears to suggest that change rather than undesirability is the characteristic of life events that should be measured for the more accurate assessment of their stressfulness (Dohrenwend, 1973; Selye, 1956). However, this evidence itself is potentially constrained by the scales designed to test the hypothesis. Other evidence, relying on qualitative data (Kleinman, 1980), lead to the opposite conclusion. While there may be no resolution to this debate, when using existing scales and techniques, one should keep in mind the theoretical foundation on which the scale was developed.

A second problem in the use of life events has been weak correlations obtained in research between events and other measures of stress. In studies of depression, for instance, correlation coefficients obtained between life events and depression range between .09 and .30, indicating that life events account for only 1% to 9% of the variance in depressive disorders. Such figures clearly appear to cast doubt on the predictive significance of specific life events in depression (Hirschfeld and Cross 1982, p. 42). Similar weak correlations have been reported in prospective studies of life events and other symptoms of stress (e.g., Grant et al., 1982; Myers, Lindenthal and Pepper, 1974).

A third major problem in the use of life event scales is the difficulty of establishing a causal sequence. According to Susser (1981), some scales combine events, many of which do not occur prior to the manifestations under study. From any one such event, such as divorce, the task of separating a stress reaction to an event from self and social selection for exposure to that event has proved difficult. A divorce, rather than precipitating the onset of disease, may itself be the indirect product of that disease as individuals become divorced and remain unmarried because they are ill. Cases may be predisposed to certain events and if such a predisposition is viewed as a latent manifestation of the dependent variable and left uncontrolled, it is not amenable to firm inference.

In addressing these problems, at a mimimum, two aspects of life events should be ascer-

tained by the investigator. First, the meaning of the event should be provided by the subject and not a priori by the investigator. Second, the context in which an event occurs should be provided. This includes the social support network of the subject, preceding and succeeding events, financial resources, social status, and so on.

Another independent variable possessing certain limitations is that of social status. Frequently, simplicity and lack of uniformity in the operationalization of status tends to confound the relationships between status and disease episodes as well as make comparison between studies relying upon different definitions of status difficult, if not impossible. "Many epidemiologic studies treat social variables as if they were biological variables and simplify them to the point where their meaning is questionable" (McQueen and Siegrist, 1982, p. 353). The operationalization of social class or socioeconomic status (SES) is a significant problem in psychosocial studies. As noted above, certain studies rely upon two indices, usually education and income, as measures of socioeconomic status. Other studies rely upon single indices such as census tract or occupation. As Kessler (1982) noted in his study of indices of SES, each index has a particular effect on stress, independently influencing emotional functioning. Furthermore, these indicators vary in importance with other social characteristics such as sex and labor force status.

The operationalization of racial and ethnic group categories presents another set of problems which must be accounted for in psychosocial studies. Race and ethnicity are often used interchangeably but the two are not isomorphic. Different ethnic groups can belong to one racial category, as Navaho and Sioux constitute two separate ethnic groups but one racial group. Conversely, the same ethnic group can include individuals belonging to different racial groups, as in the case of Black and Caucasian Puerto Ricans. Where possible, ethnic groups are to be preferred when looking at behavioral factors. Use of race such as Blacks or non-whites can mask important differences. A high rate of alcohol abuse among non-whites in Alaska, for example, might disguise the fact that such abuse is much lower for Blacks, Chinese and Vietnamese in Alaska than it is for Eskimos, Aleuts and Athapaskans. Because of the variable nature of definitions of ethnic group status, self-designations are to be preferred over the use of other characteristics such as physical features, language, surname, and place of birth.

Social support is another independent variable in psychosocial studies possessing certain limitations. According to Thoits (1982), most studies suffer from inadequate conceptualization and operationalization of social support. The direct effect of life events upon social support is often confounded, either theoretically or operationally, with the interactive (buffering) effect of events with support. As in the case of life events, there also is the problem of evaluating the quantity vs quality of support and the problem of determining whether social support is an independent or a dependent variable.

When attempting to operationalize social support, several considerations need to be taken into account. First, amount of support, types of support (e.g. socioemotional and instrumental) and sources of support (e.g., spouse, friends, kin, coworkers) available are all important dimensions of social support and measures should be employed which tap each of

these dimensions. Second, a distinction should be made between measures of social support structure and function. As Thoits (1982) asserts, the structure of the social support network may have a powerful influence on the flow of supportive resources to an individual. The structural properties of the social support system relative to the total social support network can be measured using such indicators as size, density, accessibility, kinship-reliance, frequency of contact, and stability. The functional properties of the system can also be operationalized to measure the perceived amount and adequacy of socio-emotional and instrumental aid received from various support system members.

DATA ANALYSIS

With each type of research design in psychosocial epidemiology, there are numerous techniques of data analysis which may be employed to control the effects of different variables, test hypotheses, and account for associations between disease or stress episodes and one or more characteristics. Some of these techniques such as the calculation of relative risk, may be employed in several types of studies while other techniques such as the use of logistic regression models, are specifically suited to particular research designs. Whatever the research design, however, it is always recommended that data analysis proceed from the simple to the complex. All too often, sophisticated techniques are employed when simple ones provide a much clearer picture of the phenomena under investigation. The first step in any study should always be the description of the phenomena under investigation. It is here that the use of rates may be employed. The search for associations is the second step, employing such methods as proportions, ratios, and correlation coefficients. Multivariate analyses is the third step and should be employed to control for extraneous variables and to assess their role in any causal relationship.

Measures of Association

Relative Risk The most common measure of association in epidemiologic studies is relative risk which reflects the incidence of disease among a group possessing a characteristic relative to a group without the characteristic. It is a ratio of two incidence rates and indicates the likelihood that a member of a specified population will acquire and/or succumb to a disease if he possesses the characteristic under study. Thus, a study that determines that the relative risk of schizophrenia among Black enlisted personnel in the U.S. Navy is 2.5 is stating that the risk for having an episode of schizophrenia (usually a first admission to a hospital and diagnosis) is 2.5 times greater for Blacks than Navy enlisted personnel belonging to other racial groups.

The model for the calculation of relative risk is a 2x2 table in which the number of cases and controls are compared with respect to the presence or absence of a particular characteristic (see Table 1.1). The cross products are then multiplied and divided, producing the following equation:

$$RR = \frac{ad}{bc}$$

This equation is known as an odds ratio and provides an approximation of relative risk and assumes that (1) the cases and controls have been selected at random and are representative of the larger population, and (2) the frequency of disease in a population is relatively small. If RR > 1 a positive association between the disease and the characteristic is said to exist; if RR < 1, there is a negative association; and if RR = 1, no association exists. An example of estimating relative risk from an odds ratio in a 2x2 table is found in Table 1.3.

Table 1.3

Point Prevalence Rates of Major and Minor Depression by Marital Status, New Haven Study, 1967-1976

Currently Married	Cases	Controls	Total
No	14	122	136
Yes	21	354	375

$$RR = \frac{14 \times 354}{21 \times 122} = 1.9$$

Source: Weissman and Myers, 1978.

As a rule, relative risk can be exactly determined only from a prospective study. In retrospective and cross-sectional studies, the odds ratio can serve as an approximation of relative risk. If the frequency of disease in a population is large or the approximation of RR proves to be inadequate, that is, in cases where there are multiple categories of groups -- different subgroups by age and occupation -- under study, a more accurate estimate developed by Mantel and Haenszel (1959) may be employed. The revised relative risk is calculated as follows:

$$RRmh = \frac{\Sigma \frac{ad}{N}}{\Sigma \frac{bc}{N}}$$

In addition, Mantel and Haenszel have calculated summary relative risk equations for separate subcategories of exposure. The rationale for these equations is that "...overall relative risk estimates are averages and as averages may conceal substantial variation in the magnitudes of the relative risk among subgroups" (p. 740). However, a summary estimate may also be obtained by dividing the subcategories of exposure into a series of 2x2 tables if the exposure can be placed on a gradient.

Age adjustment procedures are also important when calculating relative risk. One such procedure is to calculate relative risks for each age group in the population and then compare the age-specific risks with the overall crude risk. Table 1.4 provides an example of this procedure.

A second age adjustment procedure is the matching case method in which a sample of N diseased individuals is drawn and the characteristics of each individual noted with respect to the control factors. Subsequently, a sample of N individuals without the disease, with each individual matched on the control factors to one of the diseased individuals.

In applying such a procedure, the 2x2 table takes on a different form from that shown in Table 1.1. The cell in Table 1.5 in the upper left-hand corner contains r number of pairs in which both cases and controls possess the characteristic of interest. The marginal totals (a,b,c,d) represent the entries in the cells in Table 1.1 and the total for the

Table 1.4

Age and Sex-Specific Mortality Rates and Relative Risks for Men and Women Aged 30-69 Years, Alameda County, CA, 1965-1974

Age	Number of Respondents		Number of Deaths		Death Rate*		Relative Risk
	Men	Women	Men	Women	Men	Women	Men
30-39	673	728	16	16	2.4	2.2	1.1
40-49	729	807	36	32	4.9	4.0	1.2
50-59	501	574	68	45	13.6	7.8	1.7
60-69	326	387	91	67	27.9	17.3	1.6
TOTAL	2,229	2,496	211	160	9.5	6.4	1.5**

* Percent died or deaths per 100 population.

** $p \leq .001$

Source: Berkman and Syme, 1979.

entire table is 1/2N pairs where N represents the total number of paired individuals. The calculation of the relative risk for this table would be:

$$RR = \frac{S}{t} \quad \text{(provided } t \neq 0\text{)}$$

Table 1.5

Model of Calculation of Relative Risk for Matched Cases and Controls With and Without Characteristic

	Controls		
Cases	With Characteristic	Without Characteristic	Total
With Characteristic	r	s	a*
Without Characteristic	t	u	c*
Total	b*	d*	

* a, b, c, and d are the entries in the cells of Table 1.1

A test of whether or not the observed difference between ad and bc is due to sampling variation is provided by a chi-square test for 2x2 tables. Mantel and Haenszel have developed a chi-square formula specifically for use in testing the significance of a relative risk association. When testing for significance in a matched pairs example such as in Table 1.5, the McNemar test is employed where:

$$x^2 = \frac{(|t-s|-1)}{t+s} \quad \text{with 1 df}$$

In establishing confidence limits for the test of significance, confidence limits of the logarithm (to the base e) of a corrected relative risk are computed and the logarithmic confidence levels are then reconverted to the original scale. The addition of 0.5 to the numbers a,b,c, and d corrects for a bias that can occur with small numbers of observations. Using the log-relative risk rather than the relative risk itself simplifies calculations of standard errors necessary for computing confidence intervals.

Chi-Square Tests The chi-square has two basic uses in psychosocial epidemiologic research: testing the null hypothesis and determining the significance level of an association. One of the most basic methods of assessing risk is to compare the observed rates for a specific sample with the expected rates based on the total population. Take for instance, the annual incidence rates for hypertension among air traffic controllers and second-class airmen as reported by Cobb and Rose (1973):

Table 1.6

Annual Incidence Rates per 1,000 Men for Hypertension

Age Group year	Air Traffic Controllers	Second Class Airmen
20-24	-	-
25-29	2	1
30-34	2	1
35-39	10	1
40-44	6	-
45-49	11	4
50+	15	4

		Age <40	Age 40+
Observed (O)	28.0	16.0	12.0
Expected (E)	5.1	2.4	2.7
O/E	5.6	6.6	4.4

$x^2 = 103.0$ $p \leq .0001$

Source: Cobb and Rose, 1973.

Here, the expected rates are calculated on the basis of multiplication of the age-specific rates for second class airmen by the relevant air traffic controller population and summing across age groups. Statistical significance can then be examined by the chi-square test, assuming that the expected values represent population estimates (Cobb and Rose, 1973). Using the standard chi-square formula for observed and expected frequencies, the calculation would be as follows:

$$x^2 = \frac{(O - E)}{E} = \frac{(28.0 - 5.1)}{5.6}$$

A common fallacy in employing the chi-square test is to use the chi-square value itself as a measure of the degree or strength of an association between a disease episode and a particular characteristic. Even though chi-square is excellent as a measure of the significance of an association, it does not indicate the degree or strength of association

because it is a function both of the properties of the various cells and the total number of subjects studied. The chi-square may indicate, for example, that a relative risk of 1.25 is highly significant in a study involving 2 million subjects but insignificant in a study of 2,000 subjects. The degree of association present is really only a function of the cell proportion, which explains why relative risk and odds ratios are used as measures of the strength of association.

Correlation Coefficients There are, however, measures based on the chi-square test which do provide a measure of the degree of association between an illness and a specific characteristic. One such measure is the phi coefficient. The phi coefficient or ϕ gives a numerical value, ranging from 0 to +1 for a relationship between two variables. It is calculated by using the following formula:

$$\phi = \frac{(ad - bc)}{\sqrt{(a + b)(a + c)(b + c)(b + d)}} = \frac{(ad - bc)}{\sqrt{N_1 N_2 M_1 M_2}} = \sqrt{\frac{\chi^2}{N}}$$

The phi coefficient is similar to a Pearson product-moment correlation coefficient where the covariance (C_{xy}) of the characteristic or exposure (x) and the disease (y) is divided by the square root of the product of two variances (V_x and V_y). This equation takes the form:

$$r = \frac{C_{xy}}{\sqrt{V_x V_y}}$$

Both the phi coefficient and the product-moment correlation coefficient reflects the extent to which each variable is able to predict the other. ϕ^2 is the proportion of the variance in each variable explained by the other (Morgenstern, 1982).

Another measure is Pearson's contingency coefficient where:

$$C = \sqrt{\frac{\chi^2}{\chi^2 + N}}$$

Lifetime Risk Lifetime risk is a concept developed by Thompson and Weissman (1981) for use in cohort psychiatric epidemiologic studies. It refers to the risk for onset of a particular disorder between birth and some particular age t. In considering a single birth cohort followed from birth to age t, lifetime risk to age t (LTRt) is the proportion of the cohort that would have had onset of a particular disorder by age t if all members of the cohort lived to age t or to onset of the disorder, whichever occurred first. Expressed formally, LTRt is measured in terms of f(x) which is the rate at which onset of the disorder occurs among living unaffected individuals at exactly age x. This can be expressed in the equation:

$$f(x) = \lim_{\Delta x \to 0} \left[\frac{\Pr \text{ (onset of the disorder between x and x + } \Delta x, \text{ given that the person is living and unaffected at age x)}}{\Delta x} \right]$$

Thompson and Weissman also provide two alternative measures of risk for onset of a disorder between age 0 and age t. One is the proportion of survivors affected at age t (PSAt), referred to as 'lifetime prevalence'. It denotes the proportion of a cohort that, as of age t, has ever had the disorder. The second alternative measure is the proportion of the cohort that is affected as of age t (PCAt), based on all members of the cohort,

regardless of whether they are still alive at age t.

Attributable Risk Attributable risk is frequently used in epidemiology and measures the maximum proportion of a disease attributable to a specific characteristic or etiological factor. The measure was initially defined in terms of lung cancer and smoking as the maxium proportion of lung cancer attributable to cigarette smoking in a specified population (Levin, 1953). It is expressed as:

$$AR = \frac{b(RR-1)}{b(RR-1)+1}$$

where RR = the relative risk of lung cancer among cigarette smokers as compared to nonsmokers, and b = the proportion of the total population classified as cigarette smokers. This equation is also known as the etiologic fraction when expressed as a percentage. The equation reflects two sets of comparisons: (1) smokers and nonsmokers and (2) lung cancer attributed to cigarette smoking and lung cancer related to other causes. Since only a proportion of the total population is being examined, RR-1 is employed in the calculation of attributable risk.

Table 1.7 shows the attributable risk, as a function of relative risk (RR) of different exposures or characteristics and the proportion of exposed individuals (b) in the target population. This table indicates the proportion of cases of a disease which are attributable to a specific characteristic or exposure and which would be eliminated if the characteristic or exposure were removed. The actual number of cases of disease eliminated would depend on the size of the population and the incidence rate of the disease (Schlesselman, 1982, p.44). For example, if 50 percent of a population were exposed to a particular characteristic that increased the risk of disease only fivefold, b = .5 and RR = 5, then 67 percent of the cases would be expected to be eliminated upon removal of the factor.

Table 1.7

Attributable Risks* as a Proportion of the Relative Risk (RR) and the Proportion of the Population Exposed (b)

	Relative Risk				
Proportion of Population with Characteristic or Exposed	1.5	2	4	5	10
.01	.005	.01	.02	.04	.08
.05	.02	.05	.13	.17	.31
.10	.05	.09	.23	.29	.47
.25	.11	.20	.43	.50	.69
.50	.20	.33	.60	.67	.82
.75	.27	.43	.69	.75	.87
.90	.31	.47	.73	.78	.89
.95	.32	.49	.74	.79	.90

* When multiplied by 100, these figures become etiologic fractions.

Prospective Methods

As noted above, prospective or cohort designs are preferred in epidemiology because they are less subject to problems of response bias and data collection and provide a true measure of relative risk of exposure to various characteristics or events. Typically, such studies begin with the selection of a cohort and following the cohort from time t_0 to the end of the study period t_n. However, not all members of the cohort participates or can be observed for the same length of time. In a study of mortality, members of the cohort may die before the end of the follow-up period. In any prospective study, cohort members may drop out or decide to withdraw from participation for any number of reasons which may or may not be related to the study itself. Similarly, individuals may enter the study and join the cohort at some time after the study begins and be observed until the end of the follow-up period. In both cases, these individuals may be observed for part but not all of the follow-up period. A rate of disease or stress episode based only on those observed throughout the entire follow-up period would not be an accurate reflection of the risk for the entire cohort, but only a particular segment of that cohort. The rate, therefore, would be biased by variations in the ability and willingness of cohort members to participate throughout the entire follow-up period.

In order to provide an accurate assessment of risk for the entire cohort in a prospective study, several procedures are available. Previously discussed was the use of man-years rather than number of individuals as a means of calculating incidence rates. Another technique commonly employed in prospective studies is the use of a life table to calculate the rate of morbidity or mortality across a specified number of time intervals throughout the follow-up period. Both of these techniques insure that the information provided by cohort members who do not participate during the entire follow-up period is used in the assessment of risk for particular disease or stress episodes.

Life tables are also of use in estimating the probability of surviving a particular disease episode, such as a heart attack or diagnosis of cancer. This is often known as a survival rate. This calculation must regard time as a potential extraneous variable. Adjustment for time can be effected through estimations of survival based on the life table or maximum likelihood estimates or through time series techniques such as Fourier analysis. As we shall see below, these methods are particularly useful in psychosocial epidemiologic studies of life events and stress.

Life Tables An important methodological device used in the calculation and adjustment of rates is the life table. Life tables are usually employed in studies of mortality but may also be used in morbidity and stress episode studies as well. The life table organizes data on disease or stress episodes, provides direct estimates of the probability of developing or dying from a disease for a given time period and relative risks can be computed as the ratio of these probabilities.

There are three types of life tables: current, cohort, and follow-up. The current life table uses mortality rates observed at a given time such as a census year. The cohort

life table follows a population born in the same year over time, applying mortality rates appropriate to that cohort at each age interval. The follow-up table is used to determine rates of survival during periods of observation when individuals will withdraw from the study, either because of death, change of residence, or lack of willingness to participate. Construction of a life table for a specific follow-up period does not require that the entire cohort be observed for that entire period. The table includes information on those who withdraw before the termination of a study as well as those who enter at some point after the intiation of the study. In both instances, the information can be used in the calculation of rates.

Table 1.8 provides a sample of a follow-up life table. This table organizes the observations of mortality from cancer of the kidney conducted by Cutler and Ederer (1958) between 1946 and 1951. Of the 126 patients for whom information was recorded, only 9 were present at the beginning of the study. Nevertheless, the table incorporates the experience of all patients for varying lengths of time.

Column 1. Interval (x to x + 1). This column gives the time elapsed from the date of diagnosis in intervals of one year. For example, a patient who was diagnosed Jan 20, 1946 and died Oct. 5, 1948 died during the third year after diagnosis or in interval 2-3.

Column 2. Interval mid-point (x/2). This figure is of use when calculating the average strength of the population for the interval.

Column 3. Interval width (h). Occasionally, life tables include intervals which differ in length of time. Interval width is important for standardizing survival rates which will be evident when we discuss them below.

Column 4. Alive at beginning of interval (l_x). The first line in this column represents the number of patients in the cohort.

Column 5. Number of terminal events (d_x). The number who died during the interval.

Column 6. Lost to follow-up during interval (u_x). This column includes patients whose survival status as of the end of the study was unknown. The length of observation for each patient lost to follow-up is the time elapsed between date of entry and date last known to be alive. It is assumed that subsequent to date of last contact, the survival experience of lost cases is similar to that of cases continuing to be observed.

Column 7. Withdrawn alive during interval (w_x). This includes the number of patients known to have been alive at the close of the study.

Column 8. Number exposed to risk of dying (l_x'). This column represents the number of individuals that exposed to risk in the interval, computed as the number entering the interval minus 1/2 of those withdrawn.

$$l_x' = l_x - \frac{(u_x + w_x)}{2}$$

Column 9. Proportion dying during the interval (q_x). This is also referred to as the

probability of dying during the interval.

$$q_x = \frac{d_x}{l_x},$$

This figure can be expressed as a percentage if multiplied by 100. In a study of mortality, this is known as the death rate while in a study of morbidity, it could be viewed as an incidence rate. By comparing these rates for each interval between two or more groups, an assessment of relative risk may be obtained.

Column 10. Proportion surviving the interval (p_x). This can be referred to as the probability of surviving the interval or the survival rate. It is obtained by subtracting the proportion dying during the interval from unity:

$$p_x = 1 - q_x$$

This proportion may also be expressed as a percentage by multiplying by 100.

Survival analysis Survival analysis estimates the time interval between two events, a starting event and a terminal event. It is most commonly used to characterize the survival time of patients with severe illnesses and to study the effects of different treatments on the survival rate of such patients. In this case, the starting event would be either the time of diagnosis of disease or entry into the study, if the two events differ. The terminal event would be either the patient's death or the end of the follow-up period. Groups of patients can then be compared on the basis of one or more characteristics or treatment regimens to determine if their survival rates, measured on a curve, differ significantly.

Survival analysis, along with life tables, have already been used in psychosocial studies with great success and have the potential for being utilized in research on stressful life events. The methods can be applied to most areas of research where starting and terminal events may be defined and where the interval of time between these events is of interest. In a prospective study of life events and stress, therefore, the life event under investigation, say loss of job or divorce, may be taken as the starting event and the resulting stress in a cohort, defined in terms of hospitalization, death, or a specific psychological or biochemical measure, may be taken as the terminal event.

There are several means available for estimating survival in a population. The two most common procedures are based on the life table and and maximum likelihood estimate. From the life table model above, a number of functions characterizing the distribution of survival times in a cohort can be estimated. Three of the most common are the survivorship function, hazard function, and probability density function.

The survivorship function F(t), also known as the cumulative survival rate, represents the probability of all cases surviving to the end of each interval. The hazard function is the probability per unit of time that an individual who has survived to the beginning of an interval will die in that interval. The hazard function is the same as an age-specific

Table 1.8

Follow-up Life Table and Calculations of Survival Functions
(126 Male Connecticut Residents With Localized Kidney Cancer
Diagnosed 1946-1951 and Followed Through Dec. 31, 1951.)

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9	10)	(11)	(12)	(13)
Interval (Years after Diagnosis)	Interval Midpoint	Interval Width	Number Observed at Beginning of Interval (number	Number of Events (Died during Interval)	Lost to Follow-up during Interval	Withdrawn Alive During Interval alive)	Number Exposed to Risk of Event	Proportion Dying (Incidence/ Mortality Rate)	Proportion Surviving (Probability of Event Not Occurring)	Cumulative Proportion Surviving Through End of Interval	Probability Density Function	Hazard Function
x to x+1	m_x	h_x	l_x	d_x	u_x	w_x	l'_x	q_x	p_x	P_x	f(x)	λ_x
0-1	.5	1	126	47	4	15	116.5	.40	.60	.60	.24	.50
1-2	1.5	1	60	5	6	11	51.5	.10	.90	.54	.05	.10
2-3	2.5	1	38	2	-	15	30.5	.07	.93	.50	.03	.07
3-4	3.5	1	21	2	2	7	16.5	.12	.88	.44	.05	.13
4-5	4.5	1	10	-	-	6	7.0	.00	1.00	.44	-	.00
5-6	5.5	1	4	-	-	4	-	-	-	-	-	-

Source: Cutler and Ederer, 1958.

death rate. The probability density function is the probability per unit time of dying within a given interval.

To demonstrate how these functions are derived from a life table, let us return to the data contained in table 1.8. Columns 11, 12, and 13 contain the estimates of the survivorship function, probability density function, and hazard function respectively.

Column 11 represents the cumulative proportion of individuals surviving from diagnosis through the end of the interval (P_x). This cumulative survival rate at the end of each interval is obtained by multiplying the probabilities of survival up through the present interval.

Column 12 contains the estimate of the probability density function f(x). This function is computed as follows:

$$f(x) = \frac{P_x - P_{x+1}}{h_x} = \frac{P_x q_x}{h_x}$$

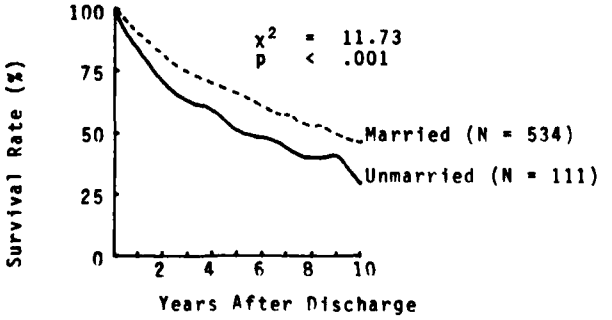
Column 13 contains the hazard rate (λ_x) which is computed as the number of deaths or episodes of the disease under investigation in the interval, divided by the average number of survivors (or those not experiencing the episode) at the mid-point of the interval.

$$\lambda_x = \frac{2(1-p_x)}{h_x(1+p_x)} = \frac{2q_x}{h_x(1+p_x)}$$

The rates obtained from these functions can be graphed in the form of curves which can be used to describe the experience of mortality from a particular disease occurring in a population. The life table procedure may also be modified for use in studies of morbidity as well, yielding similar survival rates (e.g., Chan, Powars, Lee, and Weiss, 1982). In either case, these rates can be employed in comparing the morbidity or mortality experience in two or more populations. This procedure has been employed in studies comparing survival rates for prostatic carcinoma by race and socioeconomic status (Dayal and Chiu, 1982) and for myocardial infarction by marital status (Chandra, et al., 1983). In the latter study, use of the life table to calculate survival rates among married and unmarried men yielded to following graph:

Figure 1.1

Survival Rates of Male Patients With Acute Myocardial Infarction Discharged Alive By Marital Status, Baltimore SMSA, 1966-1967 and 1971.



Source: Chandra et al., 1983

To determine whether these survival curves for married and unmarried men are significantly different, a logrank test is performed. Also known as the Mantel-Haenszel procedure for comparing the survival experience of two or more groups, this test involves a comparison between group of the observed number of events (deaths or disease episodes) with the expected number of events (Peto et al., 1976). The comparison yields a chi-square statistic which indicates level of significance. One of the advantages of this procedure is that through stratification, it allows the comparisons to be adjusted for the distribution of other factors. This becomes particularly important when examining the relations among variables of interest. Thus, in Dayal and Chiu's (1982) study of prostatic cancer, it is possible to compare the survival experience of the two races adjusting for the distribution of factors like age, stage, histologic grade, and SES and thereby conclude whether or not a particular factor 'explained' the racial difference in prostate cancer survival.

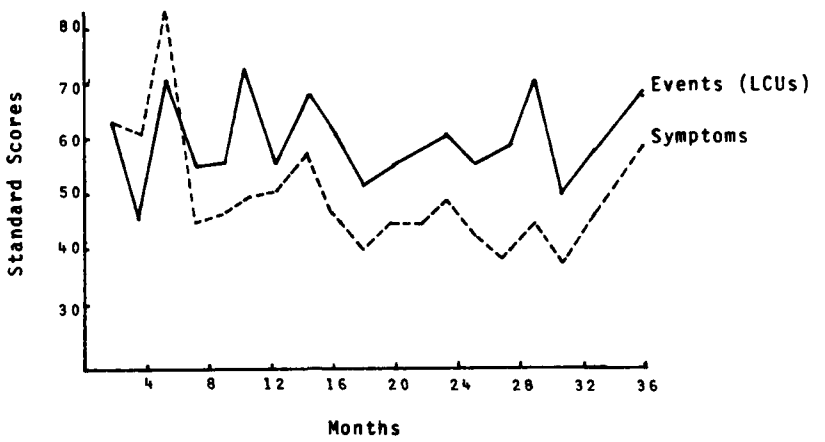
Using the life table to calculate survival functions requires grouping the data and is appropriate only when the sample size is fairly large (Gehan, 1969). A second procedure for estimating survival functions, the maximum likelihood procedure, is appropriate for small or large samples and grouped or ungrouped data. The procedure is explained in greater detail elsewhere (Kaplan and Meier, 1958).

Fourier Analysis In a prospective study by Grant, et al. (1982), Fourier analysis was used to examine possible associations between life events and psychiatric symptoms. Fourier analysis is a standard statistical technique in which time series are reduced to their sinusoidal components. In this study, the time series were symptoms and events measured in two-month intervals. Through the related statistical technique of spectral analysis, the relative strengths of these components are described as a function of frequency. The study of spectral relationships between two time series has been termed "cross-spectral analysis." The strength of the relationship of the two series at a given frequency is expressed as coherency (of that frequency). The sampling distribution of the squared coherency (which is somewhat analogous to R of regression analysis) was used to compute attained significance levels (P values) for the coherencies observed in study subjects (Grant et al 1982:599).

The following graph (Figure 1.2) summarizes the results of a Fourier analysis of life events and psychiatric symptoms for a sample of the study population characterized as being coherent and in phase. The solid line represents changes in events, measured by LCU scores, while the broken line represents fluctuations in symptoms, measured by Symptom Checklist (SCL) scores. Here, the series is coherent at higher frequency ranges (every four months) and the coherency squared of this frequency is 0.668, well above a defined criterion level of 0.35. The phase is 0.94, which shows that the two series are almost perfectly synchronized (A phase of 1.0 or 0.0 at a given frequency indicates that the two series are perfectly synchronized at that frequency. A phase of 0.5 means that one series lags behind the other by half that cycle length (Grant et al, 1982, p.599).

Figure 1.2

Life Event and Symptoms Time Series Classified as Coherent and In-Phase by Fourier Analysis



Source: Grant et al., 1982.

While time series analysis appears to hold promise as a technique for the organization and analysis of prospective data, it should be pointed out that even in this study, life events and symptoms were "coherent and in-phase" for only one-third of the cohort members, leading Grant and his associates to conclude that there was no significant correlation between life events and symptoms of psychiatric distress. However, this may be more a factor of the scale used to measure the variables than the Fourier analysis itself.

Adjustment and Controlling for Confounding Variables

Even though the object of a psychosocial study may be to examine the relationship between a disease or stress episode and one particular characteristic or form of exposure, inevitably there will be other characteristics or types of exposure which need to be accounted for. These extraneous variables potentially confound the association, wholly or partially accounting for the apparent effect of the characteristic or exposure under investigation, or masking an underlying true association. The confounding variable must satisfy both of two conditions: (1) it is a risk factor for the study disease; and (2) it is associated with the study exposure but is not a consequence of exposure (Schlesselman, 1982, p. 58). One of the best examples is the confounding effect of cigarette smoking with respect to an association between alcohol consumption and heart disease. While an apparent association may be detected between alcohol consumption and heart disease, this association exists because many individuals identified as heavy drinkers are also cigarette smokers and the real association is between smoking and heart disease.

Several techniques are available to control for extraneous variables which are potential confounders. In this section, three specific techniques are presented: specification, age-adjustment, and multivariate methods.

Specification The most elementary technique used to control for possible confounding variables is specification. This importance of specification has already been discussed and there is no need for further elaboration. By using specific rates, control of extraneous factors at a rudimentary level may be exercised in the assessment of risk. Age and sex, because of their relationships with many forms of disease and stress, are the most common variables for which specific rates are determined. Race is another common variable, although less so than age and sex.

Age-Adjustment Typically, adjustment procedures are applied in controlling for differences in age between groups because of the high association between age and the risk for particular disease episodes, particularly mortality. These procedures are limited in certain respects. The age-adjusted rate cannot provide an accurate description of a population but is only to be used as a summary of age-specific rates. The age-adjusted rates can also hide important differences between groups if the age-specific rates for the groups vary in different ways across the age groups selected. For example, a summary comparison of age-adjusted rates for depression among Blacks and Caucasians in the United States may disguise the fact that the rates are higher for young Blacks (age 15-24) than young Caucasians but higher for older Caucasians (age 65 and above) than older Blacks, this difference could be hidden by standardization.

Despite this limitation, the arguments for age adjustment are considerable. Fleiss (1981) provides three major reasons in favor of age-adjustment. First, a single summary index for a population is more easily compared with other summary indices than are entire schedules of specific rates. Second, if some age groups are comprised of small numbers of subjects, the age-specific rates for these groups may be too imprecise and unreliable for use in detailed comparisons. Third, for small populations, or for some groups of special interest, specific rates may not exist. Summary statistics allow for the inclusion of these subgroups in the overall assessment of risk in the absence of specific rates.

There are two major methods of age adjustment, direct and indirect. The direct method involves the application of the age-specific morbidity or mortality rates from the samples being compared with one another to the number in the same age groups of a standard population. For studies involving groups in the United States, the standard population is usually the population of the United States as determined in a decennial census. This procedure gives the number of episodes that can be expected in the standard population if the age-specific rates from the observed groups have prevailed in the standard population. As an example, take the data on mortality from the Alameda County study found in Table 1.9. By multiplying the mortality rates for males and females with the respective age categories of the total population of the United States in 1970, an estimate of the expected number of deaths in the total population can be obtained. The total number of expected deaths are then divided by the total standard population to obtain the age-adjusted rate. For both males and females, the age-adjusted rate is slightly higher than the crude rate.

The indirect method of age adjustment is more commonly used in epidemiologic studies

because of the often small size and lack of specific rates for study populations. It is preferable to the direct method when there are small numbers in particular age groups. "Rates used in direct adjustment would be based on these small numbers and would thus be subject to substantial sampling variation. With indirect adjustment the rates are more stable since they are based on a large standard population" (Friedman, 1980, p. 183).

Table 1.9

Age-Adjusted Death Rates for Men and Women of Alameda County, CA, 1965-1974,* Using the Direct Method of Age Adjustment

Age	Death Rate (per 100)		Standard Population 1970 U.S. Census	Expected Number of Cases	
	Men (1)	Women (2)	Men and Women (3)	Men (1) x (3)	Women (2) x (3)
30-39	2.4	2.2	22,537,287	540,894.9	495,820.3
40-49	4.9	4.0	24,096,893	1,180,747.8	963,875.7
50-59	13.6	7.8	21,077,046	2,866,478.3	1,644,009.6
60-69	27.9	17.3	15,608,409	4,354,746.1	2,700,254.8
Total Death Rate (crude)	9.5	6.4			
Total Population			83,319,635		
Total Expected Number of Deaths				8,942,867.1	5,803,960.4
Age-Adjusted Death Rate (per 100)				10.7	7.0

* Age-specific rates derived from Berkman and Syme, 1979.

To apply the indirect method, four things are required: (1) a crude rate ($r_{observed}$) for the study population, (2) a distribution across the selected age groups for that population (p_1 ,...,p_1), (3) a schedule of specific rates for the standard population (r_{s1}, ...,r_{sl}), and (4) the crude rate for the standard population ($r_{Standard}$) (Fleiss, 1981, p. 240-241). The first step in the procedure is to calculate an expected overall rate for the study population by applying the schedule of specific rates for the standard population. This takes the form of the equation:

$$r_{expected} = \Sigma r_s p$$

The indirect adjusted rate is then

$$r_{adjusted} = \frac{r_{observed}}{r_{expected}}$$

that is, the crude rate for the standard population r_s, multiplied by the ratio of the actual crude rate for the study population $r_{observed}$ to the crude rate $r_{expected}$, that would exist if the study population were subject to the standard population's schedule of rates (Fleiss, 1981).

One of the most common forms of indirect adjustment is the standardized mortality ratio. It is defined as the number of deaths, either total or cause-specific, in a given population expressed as a percentage of the number of deaths that would have been expected in that population if the age- and sex-specific rates in the standard population were appli-

cable. The statistic is calculated by using the formula:

$$\text{Standardized Mortality Ratio (SMR) (\%)} = \frac{\text{Observed number of deaths per year}}{\text{Expected number of deaths per year}} \times 100$$

The expected number of deaths per year of a particular population is calculated by using the equation

$$N = \Sigma pr ,$$

where p is the number of subjects belonging to a particular age group and r is the age-specific death rate in the standard population. The sum of these calculations, N, equals the total number of deaths per year. This figure is then compared with the observed number of deaths for the population under study.

An example of the use of the SMR is found in Table 1.10. Among male Alaskan Natives in the 25-34 age category, the standardized mortality ratio is 529, meaning that the experience of mortality for that age group is 529% of the mortality experience of males in the United States and over four times greater than the mortality experience of Non-Native males in Alaska.

Table 1.10

Standardized Mortality Ratios for Native and Non-Native Males
Alaska, 1974

	Alaska Male Population		U.S. Mortality Rate for males (per 1,000)	Number of Deaths Observed		Number of Deaths Expected		Standardized Mortality Ratio*	
Age	Native	Non Native		Native	Non Native	Native	Non Native	Native	Non Native
Under 5	3,570	13,049	25.06	32	63	89	327	35	19
5-14	8,413	28,469	.52	15	19	4	15	375	127
15-24	4,956	31,943	1.92	41	82	10	61	410	134
25-34	3,216	23,323	2.17	37	66	7	51	529	129
35-44	2,696	18,565	4.10	31	81	11	76	282	107
45-54	1,880	12,988	9.68	30	88	18	126	167	70
55-64	1,319	6,619	22.89	32	131	30	152	107	86
65-74	715	1,960	165.56	22	107	118	324	19	33
75+	460	761		33	90	76	126	43	71
TOTAL	26,955	137,677		273	727	363	1,258	75	58

Source: Population figures and number of deaths obtained from Kelso, 1977.
* Ratios are multiplied by 100.

The standardized mortality ratio is widely used in epidemiology and is often used as an estimate of relative risk. However, the size of the SMR is not always equal to the size of the relative risk and the discrepancy depends on the age of the population under study (c.f., Gaffey, 1976). Thus, a group exposed to a specific hazard may possess a constant relative risk with respect to age but the SMR may increase with age. The SMR, therefore, will produce a biased estimate of relative risk and its bias will be different with each

age group. Symons and Taulbee (1981), however, argue that the standardized mortality ratio may be a useful approximation of relative risk when three conditions are met: (1) the age-specific rates in the standard population for the cause of interest are larger than 1 per 100 per year; (2) the age intervals are not too broad, shorter intervals being necessary when mortality rates are small; and (3) the age range for the analysis is restricted, neither beginning at too early an age nor extending too far into the older ages (55+).

Other Methods As the number of variables increases, so does the problem of controlling for potential confounding. Analysis based on cross-classification of variables becomes unwieldy and impractical. The alternative is to employ more sophisticated quantitative methods, including analysis of variance and covariance, bivariate and multivariate regression models and discriminant analysis. It would be impractical to go into detail on the applications of each of these models in psychosocial epidemiology here. We will therefore limit ourselves to a general discussion of multivariate methods and provide an overview of one particular model used widely in epidemiology, multiple logistic regression.

Multiple regression is frequently employed in psychosocial studies for three different purposes. First, it provides a description of the linear dependence of a disease or form of stress on a set of characteristics or exposure of a particular population. In a study of coronary heart disease, for example, regression models describe the contribution of several characteristics of an individual, including age, sex, socioeconomic status, diet, blood pressure, cholesterol, weight, and so on, to the risk of experiencing a CHD event such as angina or myocardial infarction. Using this description, predictions can be made as to the risk of disease for an individual with a specific set of values for each characteristic (e.g., a 35-year old male, with a high-school education, earning $15,000 a year, etc.).

Second, regression models are used to control for potential confounding factors in order to evaluate the contribution of a specific variable or set of variables. In the Honolulu Heart Study of acculturation and coronary heart disease, for example, several covariables are controlled using a logistic model to isolate the independent effect of each form of acculturation.

Third, regression is used to find structural relations among sets of variables and to provide causal explanations for these relationships. In their study of coping and stress, for example, (Pearlin, Lieberman, Menaghan, and Mullan (1981), employed a path model to indicate the direction and strength of association between disruptive job events and changes in economic strain, mastery, self-esteem, and depression, with the life events at the beginning of the path and depression at the end. Such a technique indicates the role of other factors as intervening or mediating variables as well as provides a measure of the degree of presumed causal connection between independent and dependent variables.

A thorough explanation of the principles which underly regression can be found in most statistics textbooks. However, a brief outline of these principles is necessary to clari-

fy the discussion to follow. Regression is used to quantify the relationship between two or more variables when the value of one variable is affected by changes in the values of other variables. The affected variable is the dependent (y) variable and the others which are used as predictors are independent (x) variables. When there is an orderly relationship between the dependent and independent variables, information about the x value contains some information about the y value. This orderly relationship is called a correlation. If the relationship is linear, such that y is a linear function of x, the relationship may be expressed as a regression equation and the straight line which is the graph of this equation is called the regression line of y on x. In a simple linear regression where there is only one independent variable, the probability or expected value of the dependent variable (y) is expressed as:

$$p(y) = \beta_0 + \beta_1 x$$

where β_0 and β_i are the coefficients (parameters), β_0 being the intercept and β_i being the slope of the regression. The y intercept represents the predicted value of y when x = 0. The slope of the regression line β_i indicates the magnitude of the change in y for a unit change in x. A model of multiple linear regression involving k independent variables is written as

$$p(y) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_k x_k$$

where $x_1, x_2, \ldots, x_k$ are the independent variables and $\beta_1, \beta_2, \ldots, \beta_k$ are their regression coefficients.

When the dependent and independent variables are linearly related, the amount of relationship between them can be expressed by a single number called the correlation coefficient (r). With a perfect correlation such that all possible values of both the dependent and independent variables fall on a straight line, the slope will be +1.0 or -1.0 and y intercept will equal 0 and the correlation coefficient will equal +1.0 or -1.0, depending on whether the relationship is positive or negative. When the correlation is less than perfect, the points will show some degree of scatter around the regression line. As the amount of scatter increases, the value of r decreases. The regression line is that line around which the amount of scatter is smallest. The amount of scatter is most often defined in terms of the sum of squares of the vertical deviations of the points from the line. The regression line of y on x, therefore, is defined as that straight line from which the sum of the squared deviation $\Sigma(y-y')^2$ is smallest.

In order to compute the least squares line, β_0 and β_i must be computed using the formulas

$$\beta_i = \frac{N\Sigma xy - (\Sigma x)(\Sigma y)}{N\Sigma x^2 - (\Sigma x)^2}$$

$$\beta_0 = \frac{\Sigma y - \beta_i x}{N}$$

As noted above, regression models provide a means of controlling for possible confounding independent variables. A partial regression coefficient, say β_1, in a multiple regression model such as the one represented the equation

$$p(y) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_k x_k$$

stands for the expected change in y with a change in one unit of x_1, when x_2 to x_k are held constant or otherwise controlled for. Alternatively, β_1 can be seen as the expected difference in y between two groups which are different on x_1 by one unit but are the same on x_2 through x_k. Likewise, β_2 stands for the expected change in y with a unit change in x_2 when x_1 is held constant, and so on. Moreover, the combined effects of regression coefficients are additive. For example, if one were to change one unit on both x_1 and x_2, the expected change in y would be $(\beta_1 + \beta_2)$. This expected change can also be expressed as the expected difference in y between two groups which are different by one unit on both x_1 and x_2.

The relationship between the regression coefficient β and the correlation coefficient r is expressed in terms of the equation

$$\beta = r\frac{(Sy)}{(Sx)}$$

where Sy is the standard deviation of y and Sx is the standard deviation of x. When x and y are standardized to have unit variance (i.e. the standard deviations of both x and y = 1), then β = r. The standardized regression coefficients are referred to as beta weights. When two or more independent variables are measured in different units (such as body weight in pounds and blood pressure in Hg), standardized coefficients provide a useful way of assessing the relative effect of each independent variable on the dependent variable.

The significance of β can be tested either by examining the confidence interval or by evaluating the F ratio obtained from the equation

$$F = \frac{\Sigma(y' - y)/1}{\Sigma(y-y')/N-2)}$$

Several bivariate and multivariate procedures have been employed in psychosocial epidemiology for the purposes of description, controlling for confounding variables, and determining causality. The selection of a procedure depends largely on the types of questions asked in a study and the nature of the data available. Thus, data from interval scales such as blood pressure measurements or psychiatric screening scale scores may rely upon least squares regression to obtain partial correlations. Data in the form of rates might employ a binary regression analysis using percentage deviations from overall rates (Roberts and Lee, 1980). Stepwise regression techniques have been employed on variables such as coping resources to assess the affect of each variable on stress by observing the reduction in the regression of stress on other factors such as role strain as each independent variable is added into the equation (Pearlin and Schooler, 1978).

One of the most commonly used procedures in epidemiology, however, is multiple logistic analysis. This technique was developed and applied to data from the Framingham Study in an effort to identify the independent contribution of several factors to the overall risk of disease. It is more appropriate than the least squares models for analyses involving the presence or absence of disease and mortality and its regression line assumes the form of a dose-response curve rather than a straight line. In a logistic model, a dichotomized

dependent variable is predicted by one or more independent variables. One of the reasons for its wide use in epidemiology is that it provides an estimate of the relative risk of a disease associated with each variable independent of all other variables studied. Although the model is formulated in terms of prospective studies, it may also be applied to data from cross-sectional and retrospective studies.

A complete description of the logistic model and its application in epidemiology is found in Schlesselman (1982). Briefly, the model specifies that the probability of disease or stress response depends on a set of independent variables such that:

$$p(y) = p(d = 1|x) = \frac{1}{\{1 + \exp[-(\beta_0 + \beta_1 x_1 + \ldots + \beta_k x_k)]\}}$$

where d denotes either the presence (d = 1) or absence (d = 0) of the disease or stress event, and x denotes a set of k variables, $x = (x_1, x_2, \ldots, x_k)$, which may represent any potential risk factor or confounding variable, functions of them, or interactions of interest. The βs or regression coefficients are parameters representing the effects of the independent variables (characteristics or exposures) on the probability of disease or stress response. These parameters are usually derived from discriminant analysis or maximum likelihood estimation, the latter being the preferred procedure because it does not depend on any apriori assumption of multivariate normality (Schlesselman, 1982).

The logistic parameters can also be interpreted in terms of odds ratios. Using the model assumptions stated above, the risk of a disease or stress event among individuals with a particular set of characteristics or exposures with the value $x = (x_1, x_2, \ldots, x_k)$ is

$$RR = \frac{p_x}{q_x} = \exp(\beta_0 + \beta_1 x_1 + \ldots + \beta_k x_k)$$

where

$$q_x = 1 - p_x = p(d = 0|x).$$

This equation provides an approximation for the risk of disease among individuals with a specific set of values for each characteristic relative to individuals with different values for the same characteristics. An odds ratio per unit change in the level of any one characteristic, x_i is $\exp \beta_i$.

The logistic model is often expressed directly in terms of logits, $\ln(p_x/q_x)$. Thus,

$$\ln RR = \ln \frac{p_x}{q_x} = (\beta_0 + \beta_1 x_1 + \ldots + \beta_k x_k)$$

This is comparable in form to the multiple linear regression model decribed above. Expressed in terms of logits, a unit change in the variable x_i changes the logit of risk ($\ln p_x/q_x$) by the amount β_i.

Finally, the relative importance of characteristics or exposures may be compared in terms

of standardized coefficients. As in the case of linear regression, the standardized coefficient may be expressed as a multiple of the standard deviation of x such that

$$\beta'_i = (Sx_i).$$

Each standarized coefficient measures the change in the logit of risk resulting from a change of one standard deviation in the variable x_i. (Schlesselman, 1982).

Table 1.11 contains the standardized logistic coefficients from prevalence data in the study of acculturation and coronary heart disease among Japanese men in Hawaii from 1971 to 1977 (Reed et al., 1982). The effect of each variable (index of acculturation) and covariable (age, systolic blood pressure, serum cholesterol, etc.) included in the risk function is interpreted in terms of its regression coefficient β, which represents the effect of the variable adjusted for the effects of the other variables.

Table 1.11

Standardized Logistic Coefficients Relating Acculturation Scores and Selected Covariables to Coronary Heart Disease (CHD) Prevalence Among Men of Japanese Ancestry, Honolulu Heart Program, 1971-1979.

Scores	Total CHD	Myocardial infarction	Angina
Culture of upbringing	-0.08622	-0.01907	-0.15644
Current cultural assimilation	-0.09974	-0.04109	-0.15659
Current social assimilation	-0.19436**	-0.18058	-0.18232
Total Acculturation	-0.46745**	-0.70480*	-0.67075*
Covariables			
Age	0.23305**	0.13838	0.33389***
Systolic blood pressure	0.21867***	0.25565**	0.11182
Serum cholesterol	0.18533**	0.19074*	0.17142
Alcohol (ounces/month)	-0.50466***	-0.70018***	-0.20460
Complex carbohydrate (g/24 hr)	-0.22961	-0.00940	-0.23972
Serum uric acid	0.12972	0.22529*	0.03894
Serum glucose	0.10806	0.07756	0.13328
Forced vital capacity	-0.13196	-0.29020**	0.11367
Physical activity index	-0.15514*	-0.05848	-0.24166*
Cigarettes/day	0.04331	0.17335	-0.10199
Body mass index	0.08428	-0.15261	0.18531
Socioeconomic status	-0.10724	-0.01103	-0.12523

* $0.05 \geq p > 0.01$
** $0.01 \geq p > 0.001$
*** $p \leq 0.001$

Source: Reed et al., 1982.

Here it can be seen that the total acculturation score is significantly associated with total coronary heart disease as well as with myocardial infarction and angina in particular, independent of the covariables. The negative correlation indicates that respondents with a greater number of traditional Japanese responses on the acculturation scales have a lower prevalence of coronary heart disease. With the exception of current social assimilation which was associated with total CHD, none of the separate measures of acculturation were significantly associated with total or specific CHD prevalence. The use of standardized coefficients also allows for a comparison of the relative importance of each of the independent variables. Thus, age and systolic blood pressure are of equal importance in predicting the prevalence of total coronary heart disease episodes among men of Japanese ancestry in Hawaii, both covariables having standardized coefficients approxi-

mately equal to 0.3.

Associations and Causality

While epidemiologic methods have been applied with great success in detecting associations between disease or stress episodes and a host of psychosocial variables, the problem of verifying causal relations in psychosocial epidemiology has been much more complex. The literature is replete with examples of correlations and speculations about causality. Often, associations have been taken as evidence for causal relations. In other studies, statements about causal relations are limited by the failure to control for all potential confounding variables. As noted above, multivariate methods are often employed to estimate the strength and direction of associations between dependent and independent variables. Despite the efficacy of these methods, however, several considerations must be kept in mind when examining the causal relations between a disease or stress episode and characteristics or exposures. Schlesselman (1982) lists six specific observational criteria for determining causation in an epidemiologic investigation: (1) temporal sequence, (2) consistency, (3) strength of association, (4) biological gradient, (5) specificity of effect, and (6) collateral evidence and biological plausibility.

Temporal sequence is particularly important with respect to life events and social research. While it is generally assumed that life events lead to psychological distress, they may also be a consequence of distress. The causal association between migration and psychological distress is an example of this issue. While research has shown migrants to have higher levels of mental disorder than nonmigrants, cross-sectional or retrospective studies are of little use in determining whether the disorder is due to the immigration process or whether the disorder itself precipitated the migration (self-selection). To demonstrate a causal relation between life events and distress, the life events must be shown to occur prior to the onset of the disorder. In their study of psychological impairment and life events, Andrews et al. (1978), attempted to control for this temporal sequence by studying the cause of life events among cases and controls and dating events at monthly intervals, making it possible to examine the correlation between symptom score and life events in specific time periods.

The consistency of an association under different conditions of study increases the likelihood of a causal relation between a disorder and a characteristic or exposure. Demonstrating an association between loss of spouse and myocardial infarction among men from different occupations, belonging to different racial and socioeconomic status groups, and living in different geographic areas (urban vs rural) gives added weight to a hypothesized causal relationship between the characteristic and the disease.

The strength of an association as indicated by relative risk or correlation coefficients also aids in the confirmation of a causal relationship. The larger the relative risk or correlation coefficient, the less likely the association is spurious.

The existence of a biological gradient is often used in biomedical studies to indicate a

causal association. This gradient, often represented as a dose-response curve, makes a causal interpretation more plausible. In psychosocial studies, such a gradient typically involves biochemical markers or screening scale scores as dependent variables and interval scale measures of exposure.

A causal association may also be indicated by determining the specificity of the effect of one or more independent variables. A characteristic or exposure is considered to be specific to a disease or stress response if the introduction of the former is followed by the occurrence of the latter, and if the removal of the former is followed by the absence of the latter. In psychosocial studies, however, specificity is difficult to achieve and its usefulness limited by the complex interrelations of typical characteristics of interest and the disease or stress response where multiple causes and effects are usually the rule.

Finally, the ability to marshal collateral evidence aids in the determination of a causal relation between variables. Arguments for the existence of causal relationships between certain psychosocial characteristics such as crowding and competition and stress responses such as hypertension, for example, have used data from animal studies and experimental studies on humans as well as observational studies of population samples (Henry and Cassel, 1969). While the collateral data may have been collected with different objectives in mind, they may provide important clues as to the strength and direction of associations between variables.

While each of these conditions alone is often insufficient to establish causality in associations between independent and dependent variables, in combinations they increase the level of confidence with which statements of causality are made.

QUALITATIVE TECHNIQUES IN PSYCHOSOCIAL EPIDEMIOLOGY

Throughout this chapter we have been concerned primarily with quantitative techniques as they are employed in psychosocial epidemiology. However, it would be difficult to overemphasize the importance of qualitative techniques as well. Qualitative techniques such as participant observation, informal interviews, taxonomic and thematic analyses can be of critical use in all aspects of psychosocial epidemiology. This last section will briefly touch upon the necessity for a qualitative perspective in three specific stages of a psychosocial study: study design, data collection, and data analysis.

Study Design

Qualitative techniques are often employed in the initial stages of an epidemiologic investigation to "scope out" the problem at hand and to identify at a broad level the relevant study parameters. Informal interviews and observation of study populations can be used to identify the issues involved in studying a particular disease or stress response, potential problems in data collection, potential confounding variables unrecognized in previous research, and characteristics specific to the study population--i.e., how it is like or

unlike the general population and other groups in similar studies.

Qualitative techniques are also employed at the initial stages of a psychosocial study to develop measures of variables of interest. In Pearlin and Schooler's (1978) study of coping and emotional stress, for example, measures of life strains and coping responses were developed from a thematic analysis of answers obtained from unstructured, open-ended interviews. This information was then subjected to a factor analysis which identified certain constructs such as self-esteem or self-denigration and the questions most salient to measuring the construct.

Data Collection

The value of qualitative methods at this stage is determined by the type of data gathered and the methods used to collect them. Given the relationships between the social, psychological, and physiological components of stress and illness, an understanding of the problem usually requires at least some level of qualitative information for two reasons. First, behavior, including illness behavior and stress responses, is guided, at least in part, by numerous cultural meanings. In fact, as Kleinman (1980) observes, illness can be looked upon as a system of meaning, where perception, evaluation and treatment of particular symptoms will be largely influenced by the cultural pattern adhered to by the individual. Second, illness occurs in a variety of social and environmental contexts, some of which are associated with the illness itself, others which are not. While quantitative methods can serve to identify certain features of the context in which a disease or stress episode occurs, a complete description often requires qualitative data as well.

An understanding of both the context in which a disease or stress episode occurs and the meaning the event has for an individual is crucial for a number of reasons. First, both context and meaning influence the presentation and severity of symptoms and styles of coping with disease or stress. As noted above, coping is strongly influenced by several factors such as previous experience, cultural values, and personality, some of which are largely unquantifiable, others which are better understood through qualitative techniques such as life histories and clinical evaluations. Second, the context and meaning of a particular disease or stress response will determine the extent to which quantitative techniques are valid and reliable in the data collection stage. At several points in this chapter we have noted the potential problems that can occur with certain types of data. Rates of incidence can be distorted by patient preference for treatment outside the formal health care system, by response bias due to coping strategies such as denial or overreporting, and by the quality of interaction between researcher and subject. In each instance, qualitative data is invaluable in recognizing the limitations of certain types of data and in selecting quantitative techniques best suited to handle these limitations.

Data Analysis

Finally, qualitative techniques are useful in the data analysis stage of an epidemiologic study. Although an association, and perhaps a causal relationship, can be demonstrated using an array of quantitative techniques, each researcher must address the questions of

whether the conclusions make sense and are realistic. The observational criteria for determining causality help to provide answers to these questions. Qualitative data can be employed as collateral evidence in establishing the validity and reliability of results based on quantitative analyses. If a similar conclusion can be reached with a different perspective, the validity and reliability of the conclusion is strengthened.

LIST OF REFERENCES

Adebimpe, V. B. (1981) Overview: White norms and psychiatric diagnosis of Black patients. American Journal of Psychiatry, 138, 279-285.

Andrews, G., Tennant, C., Hewson, D., and Vaillant, G. (1978) Life event stress, social support, coping style, and risk of psychological impairment. Journal of Nervous and Mental Disease, 166, 307-316.

Aneshensel, C. S., and Stone, J. D. (1982) Stress and depression: a test of the buffering model of social support. Archives of General Psychiatry, 39, 1392-1396.

Antonovsky, A. and Kats, R. (1967) The life crisis history as a tool in epidemiologic research. Journal of Health and Social Behavior, 8, 15-21.

Appley, M., and Trumbull, R. (1967) On the concept of psychological stress. In M. Appley and R. Trumbull (Eds.), Psychological Stress, 1-13. New York: Appleton-Century-Crofts.

Baum, A., Grunberg, N. E., and Singer, J. E. (1982) The use of psychological and neuroendocrinological measurements in the study of stress. Health Psychology, 1, 217-236.

Berkman, L. F., and Syme, S. L. (1979) Social networks, host resistence, and mortality: a nine-year follow-up study of Alameda County residents. American Journal of Epidemiology, 109, 186-204.

Berkson, J. (1946) Limitations in the application of fourfold table analysis to hospital data. Biometrics, 2, 47-53.

Berkson, J., and Gage, R. P. (1950) Calculation of survival rates for cancer. Proceedings of the Staff Meetings of the Mayo Clinic, 25, 270-286.

Brown, D. E. (1981) General stress in anthropological fieldwork. American Anthropologist, 83, 74-92.

Brown, G. W., and Birley, J. L. (1968) Crisis and life changes and the onset of schizophrenia. Journal of Health and Social Behavior, 9, 203-214.

Byrne, D. G. and Whyte, H. M. (1980) Life events and myocardial infarction revisited: the role of measurements of individual impact. Psychosomatic Medicine, 42, 1-10.

Cannon, W. B. (1929) Bodily Changes in Pain, Hunger, Fear, and Rage: An Account of Recent Researches into the Function of Emotional Excitement. New York: Appleton-Century-Crofts.

Celentano, D. D., and McQueen, D. V. (1978) Comparison of alcoholism prevalence rates obtained by survey and indirect estimators. Journal of Studies on Alcohol, 39, 420-434.

Clark, V. A., Aneschensel, C. S., Fredrichs, R.R., et al. (1981) Analysis of effects of sex and age in response to items on the CES-D scale. Psychiatry Research, 5, 171-181.

Chan, L. S., Powars, D., and Weiss, J. (1982) A modified life table method to study congenital genetic disorders: an application in sickle cell anemia. Journal of Chronic Diseases, 35, 401-409.

Chandra, V., Szklo, M., Goldberg, R., and Tonascia, J. (1983) The impact of marital status on survival after an acute myocardial infarction: a population-based study. American Journal of Epidemiology, 117, 320-325.

Cobb, S., and Rose, R. M. (1973) Hypertension, peptic ulcer, and diabetes in air traffic controllers. Journal of the American Medical Association, 224, 489-492.

Cohen, S. (1980) Aftereffects of stress on human performance and social behavior: a review of research and theory. Psychological Bulletin, 87, 578-604.

Cutler, S. J., and Ederer, F. (1958) Maximum utilization of the life table method in analyzing survival. Journal of Chronic Diseases, 8, 699- 712.

Dayal, H., and Chiu, C. (1982) Factors associated with racial differences in survival for prostatic carcinoma. Journal of Chronic Diseases, 35, 553-560.

Dohrenwend, B. S. (1973) Life events as stressors: a methodological inquiry. Journal of Health and Social Behavior, 14, 167-175.

Dohrenwend, B. S. (1980) The conflict between statistical and theoretical significance. Journal of Health and Social Behavior, 41, 291-293.

Dohrenwend, B. P., and Dohrenwend, B. S. (1969) Social Status and Psychological Distress: A Causal Inquiry. New York: John Wiley and Sons.

Dohrenwend, B. P., and Dohrenwend, B. S. (1982) Perspectives on the past and future of psychiatric epidemiology. American Journal of Public Health, 72, 1271-1279.

Dohrenwend, B. P., Shrout, P. E., Egri, G., and Mendelsohn, F. S. (1980) Nonspecific psychological distress and other dimensions of psychopathology. Archives of General Psychiatry, 37, 1229-1236.

Feinstein, A. R. (1979) Methodological problems and standards in case-control research. Journal of Chronic Diseases, 18, 35-41.

Fleiss, J. L. (1981) Statistical Methods for Rates and Proportions. (2nd ed.) New York: John Wiley and Sons.

Fox, B. H. (1978) Premorbid psychological factors as related to cancer incidence. Journal of Behavioral Medicine, 1, 45-133.

Frankenhaeuser, M. (1975) Sympathetic-adrenomedullary activity, behavior and the psychosocial environment. In P. H. Venables and M. J. Christie (Eds.), Research in Psychophysiology, 71-94. New York: John Wiley and Sons.

Fried, M. (1969) Social differences in mental health. In J. Kosa, A. Antonovsky, and I. Zola (Eds.), Poverty and Health: A Sociological Analysis, 113-167. Cambridge: Harvard University Press.

Friedman, G. D. (1980) Primer of Epidemiology. (2nd ed.) New York: McGraw Hill.

Gaffey, W. R. (1976) A critique of the standardized mortality ratio. Journal of Occupational Medicine, 18, 157-160.

Gaines, T., and Richmond, L. H. (1980) Assessing suicidal behavior in basic military trainees. Military Medicine, 145, 263-266.

Gehan, E. A. (1969) Estimating survival functions from the life table. Journal of Chronic Diseases, 21, 629-644.

Glass, D. C. and Singer, J. E. (1972) Urban Stress: Experiments in Noise and Social Stressors. New York: Academic Press.

Grant, I., Yager, J., Sweetwood, H. L., and Olshen, R. (1982) Life events and symptoms: Fourier analysis of time series from a three-year prospective inquiry. Archives of General Psychiatry, 39, 598-605.

Hahn, M. E. (1966) California Life Goals Evaluation Schedule. Palo Alto: Western Psychological Services.

Harburg, E., Gleibermann, L., Roeper, P., Schork, M. A., and Schull, W. J. (1978) Skin color, ethnicity, and blood pressure I: Detroit Blacks. American Journal of Public Health, 68, 1177-1183.

Hayes-Bautista, D. E. (1983) On comparing studies of different Raza populations. American Journal of Public Health, 73, 274-276.

Haynes, S. G., and Feinleib, M. (1980) Women, work and coronary heart disease: prospective findings from the Framingham Heart Study. American Journal of Public Health, 70, 133-141.

Henry, J. P. (1982) The relation of social to biological processes in disease. Social Science and Medicine, 16, 369-380.

Henry, J. P., and Cassel, J. C. (1969) Psychosocial factors in essential hypertension. American Journal of Epidemiology, 90, 171-200.

Hirschfeld, R., and Cross, C. K. (1982) Epidemiology of affective disorders: psychosocial risk factors. Archives of General Psychiatry, 39, 35-46.

Holmes, T. H., and Rahe, R. H. (1967) The social readjustment rating scale. Journal of Psychosomatic Medicine, 11, 213-218.

House, J. S., Robbins, C., and Metzner, H. (1982) The association of social relationships and activities with mortality: prospective evidence from the Tecumseh community health study. American Journal of Epidemiology, 116, 123-140.

Hypertension Detection and Follow-up Program Cooperative Group (1977) Race, education and prevalence of hypertension. American Journal of Epidemiology, 106, 351-361.

Jackson, D. N. (1974) Personality Research Form Manual. Goshen, NY: Research Psychologists' Press.

Johnson, J. H., and Sarason, I. G. (1978) Life stress, depression and anxiety: internal-external control as a moderator variable. Journal of Psychosomatic Research, 22, 205-208.

Kaplan, A. (1977) Social support: the construct and its measurement. Unpublished B.A. thesis, Department of Psychology, Brown University, Providence, R.I.

Keil, J. E., Tyroler, H. A., Sandifer, S. H., and Boyle, E. (1977) Hypertension: effects of social class and racial admixture. American Journal of Public Health, 67, 634-639.

Kelso, D. (1977) Working Papers: Descriptive Analysis of the Impact of Alcoholism and Alcohol Abuse in Alaska, 1975. Vol. IV. Juneau: State Office of Alcoholism, Department of Health and Social Services, State of Alaska.

Kennedy, J. G. (1973) Cultural psychiatry, In J. Honigman (Ed.), Handbook of Social and Cultural Anthropology, 1119-1198. Chicago: Rand MacNally.

Kessler, R. C. (1979) A strategy for studying differential vulnerability to the psychological consequences of stress. Journal of Health and Social Behavior, 20, 100-108.

Kessler, R. C. (1982) A disaggregation of the relationship between socioeconomic status and psychological distress. American Sociological Review, 47, 752-764.

Kiev, A. (1972) Transcultural Psychiatry. New York: The Free Press.

Kleinman, A. (1980) Patients and Healers in the Context of Culture: An Exploration of the Borderland Between Anthropology, Medicine, and Psychiatry. Berkeley: University of California Press.

Kleinman, A., Eisenberg, L., and Good, B. (1978) Culture, illness, and care: clinical lessons from anthropologic and cross-cultural research. Annals of Internal Medicine, 88, 251-258.

Kobasa, S. C., Maddi, S. R., and Courington, S. (1981) Personality and constitution as mediators in the stress-illness relationship. Journal of Health and Social Behavior, 22, 368-378.

Langner, T. S. (1962) A twenty-two item screening score of psychiatric symptoms indicating impairment. Journal of Health and Human Behavior, 3, 269-276.

Lazarus, R. (1966) Psychological Stress and the Coping Process. New York: McGraw Hill.

Levin, M. L. (1953) The occurrence of lung cancer in man. Acta Unio International Contra Cancrum, 9, 531-541.

Lilienfeld, A. M., and Lilienfeld, D. E. (1980) Foundations of Epidemiology. (2nd ed.) New York: Oxford University Press.

Lilienfeld, D. E. (1978) Definitions of epidemiology. American Journal of Epidemiology, 107, 87-90.

Macmillan, A. M. (1957) The health opinion survey: technique for estimating prevalence of psychoneurotic and related types of disorders in communities. Psychological Reports, 3, 325-329.

Maddi, S. R., Kobasa, S. C., and Hoover, M. (1979) An alienation test. Journal of Humanistic Psychology, 19, 73-76.

Mantel, N. and Haenszel, W. (1959) Statistical aspects of the analysis of data from retrospective studies of disease. Journal of the National Cancer Institute, 22, 719-746.

Marmot, M. G., and Syme, S. L. (1976) Acculturation and coronary heart disease in Japanese-Americans. American Journal of Epidemiology, 104, 225-247.

Mason, J. W. (1968a) A review of psychoendocrine research on the pituitary-adrenal cortical system. Psychosomatic Medicine, 30, 576-607.

Mason, J. W. (1968b) A review of psychoendocrine research on the sympathetic-adrenal medullary system. Psychosomatic Medicine, 30, 631-653.

Mason, J. W. (1975) A historical review of the stress field. Journal of Human Stress, 1, 22-36.

McQueen, D. V., and Celentano, D. D. (1982) Social factors in the etiology of multiple outcomes: the case of blood pressure and alcohol consumption patterns. Social Science

and Medicine, 16, 397-418.

McQueen, D. V., and Siegrist, J. (1982) Social factors in the etiology of chronic disease: an overview. Social Science and Medicine, 16, 353-367.

Mischler, E. G., and Scotch, N. A. (1963) Sociocultural factors in the epidemiology of schizophrenia: a review. Psychiatry, 26, 315-351.

Monroe, S. M. (1982) Assessment of life events: retrospective vs concurrent strategies. Archives of General Psychiatry, 39, 606-610.

Morgenstern, H. (1982) Uses of ecological analysis in epidemiologic research. American Journal of Public Health, 72, 1336-1344.

Myers, J. K., Lindenthal, J. J., and Pepper, M. P. (1971) Life events and psychiatric impairment. Journal of Nervous and Mental Diseases, 152, 149-157.

Newcomb, M. D., Huba, G. J., and Bentler, P. M. (1981) A multidimensional assessment of stressful life events among adolescents: derivation and correlates. Journal of Health and Social Behavior, 22, 400-415.

Pearlin, L. I., Lieberman, M. A., Menaghan, E. G., and Mullan, J. T. (1981) The stress process. Journal of Health and Social Behavior, 22, 337-356.

Pearlin, L. I., and Schooler, C. (1978) The structure of coping. Journal of Health and Social Behavior, 19, 2-21.

Peto, R., Pike, M. C., Armitage, P., Breslow, N. E., Cox, D. R., Howard, S. V., Mantel, N., McPherson, K., Peto, J., and Smith, P. G. (1977) Design and analysis of randomized clinical trials requiring prolonged observation of each patient: analysis and examples. British Journal of Cancer, 35, 1-37.

Rahe, R. H., McKean, J. D., and Arthur, R. J. (1967) A longitudinal study of life-change and illness patterns. Journal of Psychosomatic Research, 10, 355-366.

Reed, D., McGee, D., Cohen, J., Yano, K., Syme, S. L., and Feinleib, M. (1982) Acculturation and coronary heart disease among Japanese men in Hawaii. American Journal of Epidemiology, 115, 894-905.

Roberts, R. E., and Lee, E. S. (1980) The health of Mexican Americans; evidence from the Human Population Laboratory studies. American Journal of Public Health, 70, 375-384.

Rotter, J. B., Seeman, M., and Liverant, S. (1962) Internal vs external locus of control of reinforcement: a major variable in behavior theory. In N. F. Washburne (Ed.), Decisions, Values and Groups, 473-516. London: Pergamon Press.

Sarason, I. G., Johnson, J. H., and Siegel, J. M. (1978) Assessing the impact of life changes: development of the Life Experiences Survey. Journal of Consulting and Clinical Psychology, 46, 932-946.

Schlesselman, J. J. (1982) Case-Control Studies: Design, Conduct, Analysis. New York: Oxford University Press.

Seeman, M., and Anderson, C. S. (1983) Alienation and alcohol: the role of work, mastery, and community in drinking behavior. American Sociological Review, 48, 60-77.

Selye, H. (1956) The Stress of Life. New York: McGraw Hill.

Susser, M. (1981) The epidemiology of life stress. Psychological Medicine, 11, 1-8.

Symons, M. J., and Taulbee, J. D. (1981) Practical considerations for approximating relative risk ny the standardized mortality ratio. Journal of Occupational Medicine, 23, 413-416.

Thoits, P. A. (1982) Conceptual, methodological, and theoretical problems in studying social support as a buffer against life stress. Journal of Health and Social Behavior, 23, 145-159.

Thompson, W. D., and Weissman, M. M. (1981) Quantifying lifetime risk of psychiatric disorder. Journal of Psychiatric Research, 16, 113-126.

Turner, R. J. (1981) Social support as a contingency in psychological well-being. Journal of Health and Social Behavior, 22, 357-367.

Vaillant, G. E. (1976) Natural history of male psychological health: V. the relation of choice of ego mechanisms of defense to adult adjustment. Archives of General Psychiatry, 33, 535-545.

Vernon, S. W., and Roberts, R. E. (1982) Use of the SADS-RDC in a tri-ethnic community survey. Archives of General Psychiatry, 39, 47-52.

Vernon, S. W., Roberts, R. E., and Lee, E. S. (1982) Response tendencies, ethnicity, and depression scores. American Journal of Epidemiology, 116, 482-495.

Warheit, G. J., Holzer, C. E., and Schwab, J. J. (1973) An analysis of social class and racial differences in depressive symptomatology: a community study. Journal of Health and Social Behavior, 14, 291-299.

Weissman, M. M. and Myers, J. K. (1978) affective disorders in a U. S. urban community. Archives of General Psychiatry, 35, 1304-1311.

Wing, J. K., Cooper, J. E., and Sartorious, N. (1974) The Measurement and Classification of Psychiatric Symptoms. New York: Cambridge University Press.

Wingard, D. L. (1982) The sex differential in mortality rates: demographic and behavioral factors. American Journal of Epidemiology, 115, 205-216.

Wyler, A.R., Masuda, M., and Holmes, T. H. (1968) Seriousness of illness rating scale. Journal of Psychosomatic Research, 11,363-375.

ACKNOWLEDGEMENTS

The author wishes to thank Dr. John S. Petterson and Dr. Bruce M. Harris for their critical reading of and editorial suggestions to this manuscript.

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER: 83-22	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle): TECHNIQUES OF PSYCHOSOCIAL EPIDEMIOLOGY		5. TYPE OF REPORT & PERIOD COVERED: FINAL
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s): Lawrence A. Palinkas, Ph.D.		8. CONTRACT OR GRANT NUMBER(s)
9. PERFORMING ORGANIZATION NAME AND ADDRESS: Naval Health Research Center, P.O. Box 85122, San Diego, CA 92138		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: MR0000.01.01-6033
11. CONTROLLING OFFICE NAME AND ADDRESS: Naval Medical Research & Development Command, Naval Medical Command, National Capital Region, Bethesda, MD 20814		12. REPORT DATE: October 1983 13. NUMBER OF PAGES: 62
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office): Commander, Naval Medical Command, Department of the Navy, Washington, DC 20372		15. SECURITY CLASS. (of this report): UNCLASSIFIED 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

Approved for public release; distribution unlimited

18. SUPPLEMENTARY NOTES

Chapter 1 in Measurement Strategies in Health Psychology, P. Karoly (ed.) New York: John Wiley & Sons.

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

psychosocial epidemiology, quantitative methods, research design, stress, morbidity, psychometric scales

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Psychosocial epidemiology is distinguished from other forms of epidemiologic research by its emphasis on the role of stress in the distribution of disease and its use of psychometric scales. This report provides a summary of the quantitative techniques employed in psychosocial studies in an attempt to provide a better conception of this subfield of epidemiology. The summary contains a review of research design, data collection, and data analysis. Prospective, retrospective and cross-sectional designs are examined and the advantages and disadvantages of each design evaluated. The data for psychosocial studies are

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE S/N 0102-LF-014-6601

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

usually represented in the form of rates and scales. The major dependent variables are typically disease and stress events while independent variables include sociodemographic characteristics, life events, social support indices, and personality measures. Methods of data analysis discussed include relative risk, chi-square tests, correlation coefficients, lifetime risk and attributable risk. Methods of association used specifically in prospective studies include life tables, survival analysis, and Fourier analysis. Multivariate analyses are used to adjust and control for potential confounding variables and test hypotheses. The last section of the report briefly examines the necessity of a qualitative perspective in each of the stages of a psychosocial study.

END

FILMED

1-84

DTIC